

OUTLIER EIGENVALUE FLUCTUATIONS OF PERTURBED IID MATRICES

ANAND B. RAJAGOPALAN

ABSTRACT. It is known that in various random matrix models, large perturbations create outlier eigenvalues which lie, asymptotically, in the complement of the support of the limiting spectral density. This paper is concerned with fluctuations of these outlier eigenvalues of iid matrices X_n under bounded rank and bounded operator norm perturbations A_n , namely with $\lambda(\frac{X_n}{\sqrt{n}} + A_n) - \lambda(A_n)$. The perturbations we consider are allowed to be of arbitrary Jordan type and have (left and right) eigenvectors satisfying a mild condition. We obtain the joint convergence of the (normalized) asymptotic fluctuations of the outlier eigenvalues in this setting with a unified approach.

1. INTRODUCTION

1.1. Background. Following the works of [3] and [4] investigating the asymptotic spectrum of perturbed empirical covariance matrices or spiked population models, various efforts have been undertaken to better understanding the outlier eigenvalues of perturbed random matrix models. In the Hermitian setting, the works of [7], [8], [16], [17], [11], and [12] build up to an essentially complete picture of the asymptotic locations and normalized fluctuations of the outlier eigenvalues of bounded rank and bounded operator norm perturbations.

This paper obtains the asymptotic fluctuations of outlier eigenvalues for the iid matrix ensemble under the same class of perturbations. Before stating our results, we introduce the theorem on the asymptotic location of the outlier eigenvalues due to [20] after presenting some introductory definitions and results.

Definition 1. A iid matrix X is an infinite array of (complex) iid random variables $(x_{i,j})_{i,j \geq 1}$ which we identify with the sequence $(X_n)_{n \geq 1}$, $X_n = (x_{i,j})_{1 \leq i,j \leq n}$. We assume that the atom distribution $x = x_{1,1}$ satisfies the moment conditions $\mathbb{E}x = 0$ and $\mathbb{E}|x|^2 = 1$. We let $\Lambda(Y)$ denote the spectrum of Y and let

$$\mu_n := \frac{1}{n} \sum_{\lambda \in \Lambda(X_n)} \delta_{\frac{\lambda}{\sqrt{n}}}$$

denote the empirical spectral distribution of X .

Theorem 1 (Circular law). For an iid matrix X , we have

$$\mu_{X_n} \Rightarrow \mu_C := \frac{1}{\pi} \mathbb{1}_{\{z \in \mathbb{C}: |z| \leq 1\}}$$

almost surely, where \Rightarrow denotes weak convergence.

The circular law, which is the work of many authors (see [21] and references therein), in particular implies that the spectral radius of X/\sqrt{n} , $\rho(X/\sqrt{n})$, satisfies $\limsup \rho(X/\sqrt{n}) \geq 1$ almost surely. The following is a complementary result; see [2] for a proof.

Theorem 2. *Let X_n be an iid matrix with atom distribution having bounded fourth moment. Then*

$$\rho\left(\frac{X}{\sqrt{n}}\right) = \lim_{l \rightarrow \infty} \left\| \left(\frac{X}{\sqrt{n}}\right)^l \right\|^{1/l}$$

converges to 1 almost surely as $n \rightarrow \infty$. Moreover, for $l \geq 1$, $\|(\frac{X}{\sqrt{n}})^l\|$ converges to $l + 1$ almost surely as $n \rightarrow \infty$.

Now let $A = A_n$ be a deterministic matrix of rank $O(1)$ and operator norm $O(1)$. We will assume for notational convenience that $\Theta = \Theta_n := \{\lambda \in \Lambda(A_n) : |\lambda| > 1\}$ is independent of n for n sufficiently large and we let m_θ denote the multiplicity of θ . Then the following theorem (due to [20], with generalizations to other models in [15], [18] and [6]) shows that outliers in the spectrum of $\frac{X}{\sqrt{n}} + A$ appear, in contrast to the situation in Theorem 2.

Theorem 3. *Let X be an iid matrix with bounded fourth moment and let A and Θ be as above. For each $\theta \in \Theta$ there exists*

$$\Lambda^\theta \subset \Lambda\left(\frac{X}{\sqrt{n}} + A\right)$$

with $|\Lambda^\theta| = m_\theta$ and for $\lambda \in \Lambda^\theta$,

$$\lambda \rightarrow \theta$$

almost surely.

To illustrate Theorem 3, in Figure 1 we have plotted the eigenvalues of a perturbed Gaussian matrix $X/\sqrt{n} + A$, with x having distribution $\mathcal{N}(0, 1)_{\mathbb{C}}$ and $n = 1000$. The two outliers near 2 correspond to the block $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and the two outliers near $1.5 + i$ are from the block $\begin{pmatrix} 1.5+i & 1 \\ 0 & 1.5+i \end{pmatrix}$ of A . Observe that the fluctuations from the Jordan block are larger; this phenomenon will be discussed later.

1.2. Model and statement of results. The focus of our paper is the fluctuations $\lambda - \theta$. More precisely, we obtain the limiting distribution of the normalized fluctuations when A is allowed to have arbitrary Jordan type and under certain sparsitiy and uniformity assumptions on the (left and right) eigenvectors of A . After introducing the main definition and theorem in this subsection, we will discuss simpler special cases in Subsection 1.3.

We now define the perturbation matrices we will consider in this paper, along with associated notation. To unify notation in this paper, for any complex vector z , we let

$$(1) \quad z^{(d)} := \begin{cases} z & : d = 0 \\ \bar{z} & : d = 1 \end{cases}$$

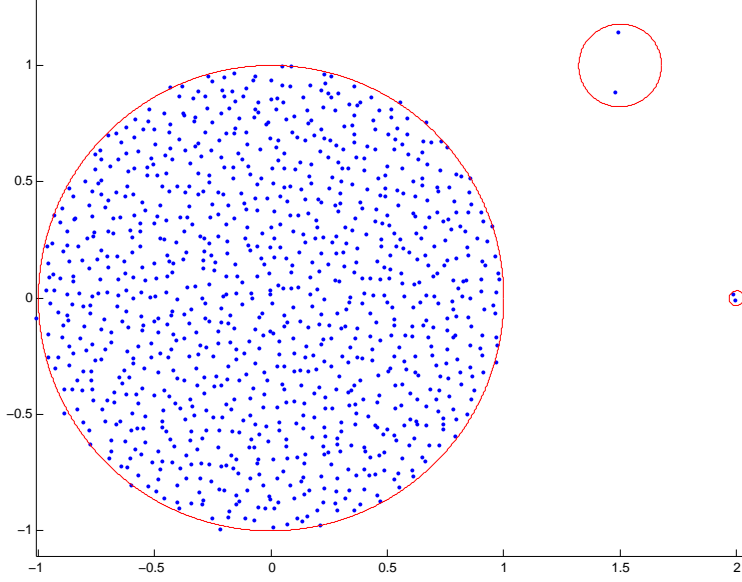


FIGURE 1. Eigenvalues of $X/\sqrt{n} + A$ with X having iid $\mathcal{N}(0, 1)_{\mathbb{C}}$ entries, $A = 2I_2 \oplus J_{1.5+i, 2} \oplus 0_{996}$ and $n = 1000$. The smaller circles are of radii $n^{-1/2}$ and $n^{-1/4}$.

where \bar{z} denotes the (componentwise) conjugate of z . We will write z^T for the transpose of z and z^* for the conjugate transpose of z .

Definition 2. A perturbation matrix $A = (A_n)_{n \geq 1}$ is a sequence of (complex) $n \times n$ matrices with rank $O(1)$ and operator norm $O(1)$. For $\theta \in \Theta = \{\theta \in \Lambda(A_n) : |\theta| > 1\}$, let J_θ be the Jordan block in the Jordan decomposition of A corresponding to θ with blocks written in nonincreasing order. We will assume that Θ and $(J_\theta)_{\theta \in \Theta}$ are independent of n for n sufficiently large. Let

$$J_\theta = \bigoplus_{k=1}^{K_\theta} J_{\theta, k}^{m_{\theta, k}} \text{ where } J_{\theta, k} := \begin{pmatrix} \theta & 1 & & \\ & \theta & 1 & \\ & & \ddots & \ddots \\ & & & \theta \end{pmatrix}$$

is the Jordan block of size k occurring with multiplicity $m_{\theta, k}$ in J_θ . To index the eigenvectors and generalized eigenvectors, we introduce the following notation. Let

$$I := \{s = (i, j, k, \theta) : i \in [k], j \in [m_{\theta, k}], k \in [K_\theta], \theta \in \Theta\}$$

and for $s \in I$, we write $s = (i_s, j_s, k_s, \theta_s)$. Let

$$I_\theta = \{s \in I : \theta_s = \theta\}.$$

For fixed j, k and θ , let $(v_s)_{i=1}^k$ be the generalized eigenvectors corresponding to the j th block of $J_{\theta, k}$, and let $v_{1, j, k}$ be the eigenvector for that block. Similarly define $(u_s^*)_{i=1}^k$ to be the generalized left eigenvectors with the $u_{(k, j, k)}^*$'s being the left eigenvectors. To index the left and right eigenvectors, we let

$$I_u^\theta := \{s \in I_\theta : i_s = k_s\}$$

and

$$I_v^\theta := \{t \in I_\theta : i_t = 1\}.$$

Finally, we let

$$I_2 := \bigcup_{\theta \in \Theta} I_u^\theta \times I_v^\theta \times \{\theta\}$$

and for $r \in I_2$, we write $r = (s_r, t_r, \theta_r)$.

For $(s_i, t_i, \theta_i) \in I_2$, $i = 1, 2$, we assume that the limits of the following inner products exist and define, for $d_1, d_2 \in \{0, 1\}$, the scalars

$$(2) \quad U_{s_1, s_2}^{(d_1), (d_2)} := \lim_{n \rightarrow \infty} (u_{s_1})^{(d_1)*} \overline{(u_{s_2})}^{(d_2)},$$

$$(3) \quad V_{t_1, t_2}^{(d_1), (d_2)} := \lim_{n \rightarrow \infty} (v_{t_1})^{(d_1)T} (v_{t_2})^{(d_2)}.$$

We also assume the following convergence and define $(G_r)_{r \in I_2}$ by

$$(4) \quad (u_{s_r}^* X v_{t_r})_{r \in I_2} \Rightarrow (G_r)_{r \in I_2}.$$

Lastly, we require the following technical assumption. Fix $\delta > 0$ and let

$$L = \bigcup_{r \in I_2} \{(i, j) \in [n]^2 : |u_{s_r, i} v_{t_r, j}| \geq n^{-1/4+\delta}\}.$$

Then we assume

$$(5) \quad \left(\sum_{(i, j) \in L} u_{s_r, i} x_{ij} v_{t_r, j} \right)_{r \in I_2} \Rightarrow (G_r^L)_{r \in I_2}.$$

Remark 1. The eigenvectors satisfying the convergence criteria of (2)-(5) are quite general, and are allowed to be of local, delocal and mixed types (see Remark 4). These eigenvector requirements are similar to those of [11] and [12].

We denote the Schur complement of A in the block matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ by

$$\text{SC}(A, \begin{pmatrix} A & B \\ C & D \end{pmatrix}) := D - CA^{-1}B.$$

Recalling the notation of Theorem 3, we denote the elements of Λ^θ by λ_s^θ for $s \in I_\theta$. We now state our main theorem.

Theorem 4. Let X be an iid matrix and A a perturbation matrix. We will assume the moment hypothesis $\mathbb{E}|x|^m < \infty$, with m defined as follows. First define c through

$$(6) \quad c = \sup\{c' \geq 0 : \max_{i \in [p]} \|u_i\|_\infty \|v_i\|_\infty \ll n^{-c'}\}.$$

Then fix $\epsilon > 0$ and set

$$(7) \quad m = \min(\max(2/c, 4), 8) + \epsilon.$$

Recalling (4), (2) and (3), we define the random variables $(F_r)_{r \in I_2}$ by

$$(8) \quad F_r := G_r + g_r,$$

where $(g_r)_{r \in I_2}$ is a collection of centered complex Gaussians independent of $(G_r)_{r \in I_2}$ with mixed second moments specified by

$$(9) \quad \mathbb{E} g_{r_1}^{(d_1)} g_{r_2}^{(d_2)} = \frac{(\mathbb{E} x^{(d_1)} x^{(d_2)})^2}{\theta_{r_1} \theta_{r_2} - \mathbb{E} x^{(d_1)} x^{(d_2)}} U_{s_{r_1}, s_{r_2}}^{(d_1), (d_2)} V_{t_{r_1}, t_{r_2}}^{(d_1), (d_2)}.$$

For $\theta \in \Theta$, let $F^\theta := (F_r)_{\theta, r=\theta}$ be the $I_u^\theta \times I_v^\theta$ matrix of random variables and for $k \in [K_\theta]$, let

$$F^{\theta, k} := \text{SC}(F^\theta|_{\{(s,t): k_s, k_t \geq k+1\}}, F^\theta|_{\{(s,t): k_s, k_t \geq k\}})$$

be the $m_{\theta, k} \times m_{\theta, k}$ matrix that is the Schur complement of the indicated submatrices of F^θ . Denote the eigenvalues of $F^{\theta, k}$ by $(\tilde{\lambda}_{j,k}^\theta)_{j=1}^{m_{\theta, k}}$ whose k th roots we denote

$$(10) \quad \tilde{f}_{i,j,k}^\theta := (\zeta_k^i (\tilde{\lambda}_{j,k}^\theta)^{1/k})_{(i,j,k) \in I_\theta, \theta \in \Theta}$$

where $\zeta_k = e^{\frac{2\pi\sqrt{-1}}{k}}$. Then for each $\theta \in \Theta$, we can label the eigenvalues in Λ^θ as $(\lambda_{i,j,k}^\theta)_{(i,j,k) \in I_\theta}$ such that the normalized outlier fluctuations

$$(11) \quad f_{i,j,k}^\theta := n^{1/(2k)} \left(\lambda_{i,j,k}^\theta \left(\frac{X}{\sqrt{n}} + A \right) - \theta \right)$$

converge to $(\tilde{f}_{i,j,k}^\theta)_{\theta \in \Theta, (i,j,k) \in I_\theta}$ in the following sense. Define the subgroup S of the permutation group S_I by

$$S := \{ \pi \in S_I : \pi(s)_\theta = s_\theta, \pi(s)_k = s_k \text{ and } \pi(s)_j = \pi(t)_j \Leftrightarrow s_j = t_j \text{ for all } s, t \in I \}.$$

Let $\text{BC}(\mathbb{C}^I)^S$ denote the set of bounded continuous functions on \mathbb{C}^I invariant under the action of S . Then for $f \in \text{BC}(\mathbb{C}^I)^S$, and writing $(\tilde{f}_l)_{l \in I}$ for (10) and $(f_l)_{l \in I}$ for (11),

$$\int f d\mu_{(f_l)_{l \in I}} \rightarrow \int f d\mu_{(\tilde{f}_l)_{l \in I}}.$$

Remark 2. The moment hypothesis we require seems to be a technical limitation of the moment method that we have employed. While we need at most $8 + \epsilon$ moments in all cases, we conjecture that 4 moments always suffice. In the delocal case with $c = 1$ (i.e., $\|u_i\|_\infty, \|v_i\|_\infty \ll 1/\sqrt{n}$), we require $4 + \epsilon$ moments which almost matches the conjectured optimal. On the other hand, under the assumption of 4 moments, [6] obtains the fluctuations of certain types of local matrices (with $c = 0$) as described in the next subsection.

1.3. Discussion and related works. We now provide examples of different types of behavior for the fluctuations that illustrate Theorem 4. The first two examples are of rank 1 fluctuations.

- (i) If A has a single non zero entry θ in the top left with $|\theta| > 1$, the limiting normalized fluctuation of the outlier is the law of $x + g$ where x is the atom distribution and g is a centered complex Gaussian with $\mathbb{E}g^2 = 0$ and $\mathbb{E}|g|^2 = \mathcal{N}(0, \frac{1}{|\theta|^2 - 1})_{\mathbb{C}}$. In Figure 2, we demonstrate this non-universality in the case $\theta = 2$ and x as specified in the captions.
- (ii) If $A = \theta v u^*$ is of rank 1 with $|\theta| > 1$ and $\|u\|_\infty \|v\|_\infty = o(1)$, then the normalized fluctuation $\sqrt{n}(\lambda - \theta)$ converges to the law of a centered complex Gaussian g_θ with

$$\mathbb{E}g_\theta^2 = \frac{|\theta|^2 \mathbb{E}x^2}{|\theta|^2 - \mathbb{E}x^2} \lim_{n \rightarrow \infty} u^* \bar{u} v^T v$$

and

$$\mathbb{E}|g_\theta|^2 = \frac{|\theta|^2}{|\theta|^2 - 1} \lim_{n \rightarrow \infty} u^* u v^* v.$$

In particular, if $\mathbb{E}x^2 = 0$ and A is normal (thus u and v are unit vectors), then g_θ is a circularly symmetric Gaussian with variance $\frac{|\theta|^2}{|\theta|^2 - 1}$.

- (iii) Suppose $A = UDU^*$ is normal of rank k , with $\|u_i\|_\infty = o(1)$ for $i = 1, 2, \dots, k$. For a fixed eigenvalue $\theta \in \Theta$ of multiplicity m , the covariance formula (9) reduces to

$$\mathbb{E}g_{ab}g_{cd} = \frac{\theta^2 \mathbb{E}x^2}{\theta^2 - \mathbb{E}x^2} \lim_{n \rightarrow \infty} u_a^* \overline{u_c} u_b^T u_d$$

and

$$\begin{aligned} \mathbb{E}g_{ab}\overline{g_{cd}} &= \frac{\theta^2}{\theta^2 - 1} \lim_{n \rightarrow \infty} u_a^* u_c \overline{u_b^* u_d} \\ &= \frac{\theta^2}{\theta^2 - 1} \delta_{ac} \delta_{bd} \end{aligned}$$

Note that fluctuations of different eigenvalues are still correlated in general. We obtain asymptotically independent fluctuations for distinct eigenvalues in the following cases.

- (a) If A is real, $u_a^* \overline{u_c} = \delta_{ac}$, $u_b^T u_d = \delta_{bd}$ and the entries of $F^\theta = (g_{ab})_{a,b=1}^m$ are independent Gaussians. Depending on the Jordan structure J_θ , the normalized fluctuations converge to the appropriate roots of eigenvalues of Schur complements of submatrices of F^θ as specified in Theorem 4.
- (b) If $\mathbb{E}x^2 = 0$, $F^\theta = (g_{ab})_{a,b=1}^m$ is a scaled complex Ginibre ensemble with atom distribution g satisfying $\mathbb{E}g = 0$, $\mathbb{E}g^2 = 0$ and $\mathbb{E}|g|^2 = \frac{|\theta|^2}{|\theta|^2 - 1}$. If we now suppose further that $J_\theta = \theta I_m$, then the m fluctuations associated to θ are given by the eigenvalues of the complex Ginibre ensemble specified above. By the circular law, they lie approximately uniformly in a disk of radius $\frac{|\theta|}{(|\theta|^2 - 1)^{\frac{1}{2}}}$ for m large.
- (c) So far, the fluctuations have been of order $O(\frac{1}{\sqrt{n}})$. Suppose again that $\mathbb{E}x^2 = 0$ but that J_θ is a single Jordan block of size m . Then as remarked below Proposition 2, the m fluctuations scaled by $n^{1/(2m)}$ are given by $(e^{2\pi i j/m} g_\theta^{1/m})_{j=0}^{m-1}$ where $g_\theta = (F^\theta)_{m1}$ is the lower left entry of F^θ . Hence the fluctuations are distributed uniformly around a circle of radius $n^{-1/(2r)} g_\theta^{1/m}$. This dependence of the rate of convergence on the size of the Jordan block is illustrated by the outliers in Figure 1.

In [18], the outlier eigenvalues of perturbations of the single ring model are studied and their locations and limiting fluctuations are obtained ([18, Theorem 2.9]) for finite rank and finite operator norm perturbations of arbitrary Jordan type. Note that the special case of the Ginibre ensemble, which is an iid matrix, is contained in this model as well. Our approach to dealing with perturbations of various Jordan types is similar and relies on a deterministic perturbation result known as the Lidskii-Vishik-Lyusternik perturbation theorem (see [13], [22], [14] and references therein) which we have reproduced in Appendix A.

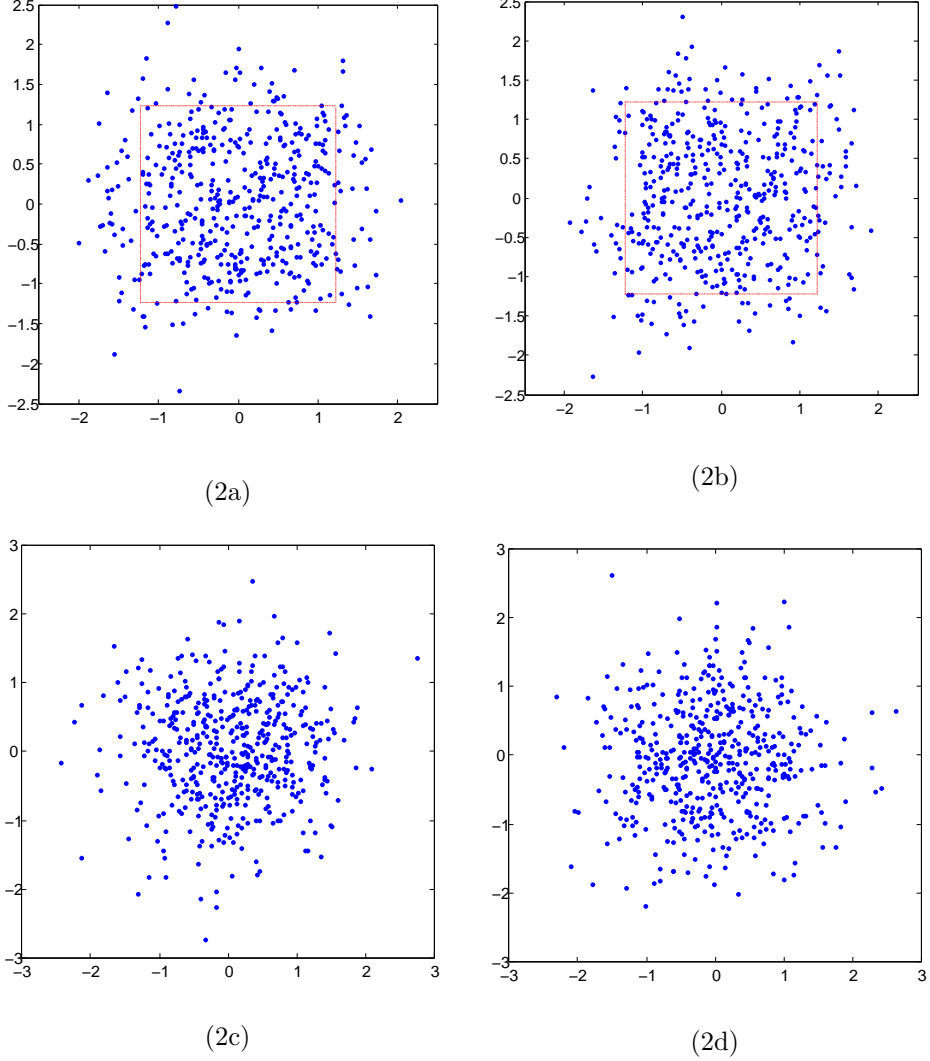


FIGURE 2. Figures 2a and 2c are 500 samples of the normalized fluctuations $\sqrt{n}(\lambda_{out}(\frac{X}{\sqrt{n}} + A) - 2)$ of a single outlier with $n = 100$ and A given by $a_{i,j} = 2\delta_{(i,j)=(1,1)}$. In Figure 2a, the atom distribution x is distributed uniformly over the square $[-l, l]^2 \subset \mathbb{C}$ with $l = \sqrt{3/2}$ so that $\mathbb{E}|x|^2 = 1$ (outlined in figure). In Figure 2c, x is the standard complex normal $\mathcal{N}(0, 1)_{\mathbb{C}}$. Figures 2b and 2d are 500 samples from the corresponding limiting distributions as predicted by Theorem 4 and detailed in case (i).

In [6], Bordenave and Capitaine study asymptotic outlier locations and fluctuations for perturbed iid matrices. The perturbations considered there are of the form $A = A' + A''$ where A'' is of bounded rank and A' (with possibly unbounded rank) satisfies a well-conditioning property. In the case of local perturbations, where A

has a finite nonzero block A'' at the top-left, [6, Theorems 1.7 and 1.8] obtain the limiting normalized outlier fluctuation when $A'' = \theta 1_{\text{rk}(A)}$ and when $A'' = J_{\theta, \text{rk}(A)}$ under the hypothesis of bounded fourth moments.

In the case when $A'' = vu^*$ is of rank 1 and is delocalized ($\|u\|_\infty, \|v\|_\infty = O(n^{-1/2})$), they show that the outliers exhibit macroscopic fluctuations and demonstrate a convergence of these fluctuations to the zeros of a Gaussian analytic function. While this phenomenon does not occur with finite rank perturbations, some techniques of the proof are similar to the ones in our proof.

In the setting of finite rank perturbations of iid matrices, when Theorem 4 is specialized appropriately, our results coincide with [18, Theorem 2.9] for the Ginibre ensemble and with [6, Theorems 1.7 and 1.8] for local perturbations of the specified Jordan types. All other cases however, with X having a non-Gaussian atom distribution and A having general eigenvectors (see Remark 1), including the delocalized cases of (ii) and (iii), do not appear to have been explicitly addressed in the literature.

The main technical result of this paper is Proposition 1 which we prove using the moment method. We require a bounded number of moments in all cases and are able to obtain the limiting fluctuations in a more general setting with a unified approach.

The paper is organized as follows. In Section 2 we prove Proposition 1 which characterizes the joint asymptotic distribution of certain random variables arising from powers of X_n appearing in the Neumann series of $(X_n/\sqrt{n} - \lambda)^{-1}$. In Section 3 we prove Lemma 7, which determines the joint limiting distribution of random variables related to a normalized resolvent of X_n , namely of the form $\sqrt{n}u^*[(X_n/\sqrt{n} - \lambda)^{-1} + \lambda^{-1}]v$. Using Lemma 7, Theorem 4 is proven in Section 4, with the help of Proposition 2 from Appendix A, a deterministic perturbation result needed to understand the effect of Jordan blocks in perturbations. Appendix B presents the truncation argument that allows us to assume stronger hypotheses in Proposition 1 and Lemma 7.

1.4. Acknowledgments. I am indebted to my advisor, Terence Tao, for his constant guidance, support and feedback throughout the course of this work.

1.5. Notation. In this paper, n will be a parameter going to infinity and many quantities will be implicitly understood to depend on n . We will use the asymptotic notation $X = O(Y)$ and $X \ll Y$ to mean there is a constant C independent of n , but possibly dependent on other parameters, such that $X \leq CY$ for sufficiently large n . Similarly, we write $X = \Omega(Y)$ to mean for some C and sufficiently large n , $X \geq CY$. We write $X = o(Y)$ to mean $\lim_{n \rightarrow \infty} X/Y \rightarrow 0$. For a sequence of events $E = E_n$, we say E occurs with high probability (w.h.p.) if $\mathbb{P}(E_n) = 1 - o(1)$ and with overwhelming probability if $1 - \mathbb{P}(E_n) \ll n^{-c}$ for all $c > 0$. We will use \Rightarrow to denote convergence in distribution (and occasionally to denote implication) and finally, we write $[k]$ for $\{1, 2, \dots, k\}$.

2. A CENTRAL LIMIT THEOREM 1

To obtain the limiting fluctuations of the outliers in Theorem 4, we will have to derive the joint asymptotic distributions for certain bilinear averages of the recentered and normalized resolvent, namely for

$$(12) \quad S_\lambda^{u,v} := -\lambda\sqrt{n}u^*((X/\sqrt{n} - \lambda)^{-1} + \lambda^{-1})v,$$

with u and v ranging over the generalized eigenvectors of the perturbation matrix A . To this end, in this section we prove Proposition 1 which obtains the limiting joint distribution for a bounded number of terms of the Neumann series of (12). In Lemma 7 we will control the tail of (12), thus obtaining its limiting distribution.

Recall the notation introduced in (1) which we reproduce here for convenience. For any complex vector z , we let

$$z^{(d)} := \begin{cases} z & : d = 0 \\ \bar{z} & : d = 1 \end{cases}.$$

For $S \subset [n] \times [n]$, we define $X_S = (X_{ij}^S)$ through

$$X_{ij}^S = \delta_{(i,j) \in S} x_{ij}.$$

Proposition 1. *Let X be an iid matrix and $(u_i, v_i)_{i=1}^p = (u_i^{(n)}, v_i^{(n)})_{i=1}^p$ be a sequence of vectors in \mathbb{C}^n . We assume the hypotheses of Theorem 4 with $(u_i, v_i)_{i=1}^p$ in the place of $(u_{s_r}, v_{t_r})_{r \in I_2}$. Thus, in the place of (2) and (3), we assume the following limits and define the scalars*

$$(13) \quad C_{i_1, i_2}^{(d_1), (d_2)} := \lim_{n \rightarrow \infty} (u_{i_1})^{(d_1)*} \overline{(u_{i_2})^{(d_2)}} (v_{i_1})^{(d_1)T} (v_{i_2})^{(d_2)}.$$

We will assume $\mathbb{E}|x|^m < \infty$ with m defined via (6) and (7).

Define

$$Z_{i,j} = Z_{i,j}^{(n)} := \sqrt{n}u_i^* \left(\frac{X_n}{\sqrt{n}} \right)^j v_i$$

where we have suppressed the n dependence for X_n , $u_i^{(n)}$ and $v_i^{(n)}$. Also, for

$$L := \bigcup_{i \in [p]} \{(k, l) \in [n] \times [n] : |u_{i,k}v_{i,l}| \geq n^{-1/4+\delta}\}$$

and $L^c := ([n] \times [n]) \setminus L$, define

$$Z_{i,j}^L := \sqrt{n}u_i^* \left(\frac{X_L}{\sqrt{n}} \right)^j v_i$$

and

$$Z_{i,j}^{L^c} := \sqrt{n}u_i^* \left(\frac{X_{L^c}}{\sqrt{n}} \right)^j v_i.$$

For $j = 1$, we will assume that the following joint convergences in distribution and define the independent families $(G_{i,1}^L)_{i=1}^p$ and $(G_{i,1}^{L^c})_{i=1}^p$ through

$$(Z_{i,1}^L)_{i=1}^p \Rightarrow (G_{i,1}^L)_{i=1}^p$$

and

$$(Z_{i,1}^{L^c})_{i=1}^p \Rightarrow (G_{i,1}^{L^c})_{i=1}^p.$$

Also define $G_{i,1} := G_{i,1}^L + G_{i,1}^{L^c}$ so that

$$(14) \quad (Z_{i,1})_{i=1}^p \Rightarrow (G_{i,1})_{i=1}^p.$$

Then for any fixed $m \geq 1$, the pm random variables $(Z_{i,j})_{i=1,j=1}^{p,m}$ converge jointly in distribution to the law of random variables $(G_{i,j})_{i=1,j=1}^{p,m}$ with $(G_{i,j})_{i=1,j=2}^{p,m}$ specified by

- (i) The $G_{i,j}$'s are centered complex Gaussians for $j \geq 2$ with mixed second moments given by

$$(15) \quad \mathbb{E}G_{i_1,j}^{(d_1)} G_{i_2,k}^{(d_2)} = \delta_{jk} (\mathbb{E}x^{(d_1)} x^{(d_2)})^j C_{i_1,i_2}^{(d_1),(d_2)}.$$

- (ii) The collections of random variables $(G_{i,1})_{i=1}^p$ and $(G_{i,j})_{i=1,j=2}^{p,m}$ are independent.

Note in particular that for $j \neq k$, $Z_{i_1,j}$ and $Z_{i_2,k}$ are asymptotically independent.

Remark 3. We note that the case $p = 1$ and $c = 1$ is a generalization of [20, Section 4] to the complex case with weaker moment assumptions, and is a special case of [6, Theorems 6.3, 6.4].

Remark 4. The assumption of the joint convergence of $Z_{i,1}^L$ and $Z_{i,1}^{L^c}$ is satisfied under various conditions. We describe some of these below.

- (i) If each u_i and v_i have finite support in $[C]$ independent of n , we have the case of a local perturbation and the $G_{i,1}$'s are finite linear combinations of the $x_{i,j}$'s.
- (ii) If each u_i and v_i is uniformly delocalized in the sense that $\|u_i\|_\infty = o(1)$ and $\|v_i\|_\infty = o(1)$ for $i \in [p]$, then by the classical central limit theorem, the $G_{i,1}$'s are joint centered complex Gaussians with mixed second moments given by

$$\mathbb{E}G_{i_1,1}^{(d_1)} G_{i_2,1}^{(d_2)} = \mathbb{E}x^{(d_1)} x^{(d_2)} C_{i_1,i_2}^{(d_1),(d_2)}.$$

- (iii) Each u_i and v_j can be allowed to have a local and a uniformly delocalized part. Namely, we suppose that for some C independent of n and all $i \in [p]$, $\sup_{i>C} |u_i|, \sup_{i>C} |v_i| = o(1)$. In this case, the $G_{i,1}$'s are a sum of a finite linear combination of the $x_{i,j}$'s and an independent Gaussian.
- (iv) Finally, we mention an example that is not contained in the above cases. Let $p = 1$, fix $0 < r < 1$ and set $u_{1,k} = v_{1,k} = r^k c_n$ with c_n chosen such that $u^* v = \theta := 2$ say. Then $G_{1,1}$ is an infinite linear combination of the $x_{i,j}$'s with exponentially decreasing entries.

2.1. Proof of Proposition 1. Instead of assuming (7), via a truncation argument presented in Appendix B, it suffices to prove Proposition 1 under the stronger assumption that the atom distribution x satisfies the bound $|x| \leq K := o(n^M)$ with $M = 2/m$ given by

$$(16) \quad M = \max(\min(c, 1/2), 1/4) - \epsilon,$$

with c defined by (6). Furthermore, by decreasing c slightly (and decreasing ϵ), we may assume

$$\max_{i \in [p]} \|u_i\|_\infty \|v_i\|_\infty \ll n^{-c}$$

instead. We will also assume without loss of generality that $(u_i, v_i)_{i \in [p]}$ are unit vectors.

In step 1, we show that $(Z_{i,1}^L)_{i=1}^p$ is asymptotically independent of

$$(Z_{i,1}^{L^c})_{i=1}^p \cup (Z_{i,j})_{i=1,j=2}^{p,m}.$$

In step 2, we derive the joint asymptotic distribution of $(Z_{i,1}^{L^c})_{i=1}^p \cup (Z_{i,j})_{i=1,j=2}^{p,m}$. A key part of the proof is contained in Lemma 5, whose proof we postpone to the end of this section.

Step 2 employs the moment method which, together with the truncation method (see Appendix B), contributes to the moment hypothesis. The moment hypothesis decays when the random variables $(Z_{i,1}^L)_{i=1}^p$ are dealt with using the moment method; thus we deal with them separately.

We will need

Lemma 1. *Let $A^{(n)} = (A_1^{(n)}, \dots, A_k^{(n)})$, $B^{(n)} = (B_1^{(n)}, \dots, B_k^{(n)})$ and $C^{(n)} = (C_1^{(n)}, \dots, C_l^{(n)})$ be sequences of complex vector valued random variables such that*

$$(A^{(n)}, C^{(n)}) \Rightarrow (A, C) \text{ and } B^{(n)} \rightarrow_P 0.$$

Then $(A^{(n)} + B^{(n)}, C^{(n)}) \Rightarrow (A, C)$. In particular, if $A^{(n)}$ and $C^{(n)}$ are independent, then $A^{(n)} + B^{(n)}$ and $C^{(n)}$ are asymptotically independent.

Proof. This follows from the Cramér-Wold device (see [5, Chapter 1.7]) and appears in [5, Exercise 1.4.2]. \square

2.1.1. Step 1. For $j \geq 2$, define

$$Z'_{i,j} := n^{-(j-1)/2} u_i^* (X - X_L)^j v_i.$$

Note that Z_{i_1,j_1}^L and Z'_{i_2,j_2} are functions of disjoint subsets of $\{x_{rs} : r, s \in [n]\}$ and hence, $(Z_{i,j})_{(i,j) \in D}$ and $(Z'_{i,j})_{(i,j) \in D^c}$ are independent. Z_{i_1,j_1}^L and $Z_{i_2,j_2}^{L^c}$ are independent for the same reason.

By Lemma 1, it suffices to show that

$$(17) \quad E = E_{i,j} := n^{-(j-1)/2} u_i^* (X^j - (X - X_L)^j) v_i \rightarrow_P 0$$

for $i \in [p]$ and $2 \leq j \leq m$. We will need the following result.

Lemma 2. *Let u and v be unit vectors in \mathbb{C}^n and X be an iid random matrix with atom distribution having mean 0, variance 1 and bounded fourth moment. Then*

$$(18) \quad \mathbb{E} \left| u^* \left(\frac{1}{\sqrt{n}} X \right)^k v \right|^2 = O \left(\frac{1}{n} \right)$$

for any fixed $k \geq 1$.

Remark 5. *Lemma 2 is a special case of Lemma 9 which establishes the same statement for k that is allowed to grow polynomially with n . We postpone the proof to Subsection 3.1, where the result is needed in full generality. We remark that Lemma 2 can also be found in [20, Lemma 2.3].*

Fix $j \geq 2$ and let $\delta_n = \log n$ (any slowly growing function of n will suffice). By Lemma 2 and Markov's inequality, for any $k \geq 1$,

$$(19) \quad \bigcap_{m=1}^M \left\{ u_m^* \left(\frac{1}{\sqrt{n}} X \right)^k v_m \leq \frac{\delta_n}{\sqrt{n}} \right\}$$

occurs with high probability for any finite set of $2M$ unit vectors $(u_m)_{m=1}^M$ and $(v_m)_{m=1}^M$.

Recall that

$$L := \bigcup_{i \in [p]} \{(k, l) \in [n] \times [n] : |u_{i,k} v_{i,l}| \geq n^{-1/4+\delta}\}$$

where $\delta > 0$ is fixed. Since $|u_i|_2 = |v_i|_2 = 1$, we have $|L| \ll n^{1/2-2\delta}$. To control, $\|X_L\|$, we will need

Lemma 3. *Suppose $S \subset A \times B$ with $\max(|A|, |B|) \leq m$. Then $\|X_S\| \leq O(\log n \sqrt{m})$ w.h.p.*

Proof. Since $\|X_S\|$ is unchanged when restricting X_S to an $m \times m$ submatrix containing S , we may assume $m = n$. If $S = \emptyset$, Lemma 3 is a consequence of Theorem 2. Writing $X' := X_L - X_{L^c}$, we have

$$\|X_L\| \leq \frac{1}{2}(\|X\| + \|X'\|)$$

from the triangle inequality. If the atom distribution x is symmetric, applying Theorem 2 to X and X' yields the desired bound. To prove the lemma for general x , we will need a symmetrization argument from [19, Section 2.3.2] that we reproduce here for convenience. Letting X'' be an independent copy of X' , we have

$$\mathbb{E}[X' - X'' | X'] = X'.$$

Since the operator norm is a convex function, we may apply Jensen's inequality to get

$$\|X'\| \leq \mathbb{E}[\|X' - X''\| | X'].$$

Removing the conditioning on X' , we have

$$\mathbb{E}\|X'\| \leq \mathbb{E}\|X' - X''\|.$$

Now $X' - X''$ has iid entries, so applying Theorem 2, we have

$$\begin{aligned}\mathbb{P}[\|X'\| \geq \log n \sqrt{n}] &\leq \frac{\mathbb{E}\|X'\|}{\log n \sqrt{n}} \\ &\leq \frac{\mathbb{E}\|X' - X''\|}{\log n \sqrt{n}} \\ &= o(1).\end{aligned}$$

□

Applying Lemma 3 with $m = n^{1/2-2\delta}$ gives

$$(20) \quad \|X_L\| \ll (\log n) n^{1/4-\delta} \text{ w.h.p.}$$

Now let

$$X^a := \begin{cases} X & : a = 0 \\ X_L & : a = 1 \end{cases}$$

Expanding (17), we have

$$\begin{aligned}|E| &\leq \sum_{a=1}^j \sum_{\substack{a_1, \dots, a_j \in \{0,1\} \\ \sum a_i = a}} n^{-(j-1)/2} |u_i^* X^{a_1} \dots X^{a_j} v_i| \\ &=: \sum_{a=1}^k E_a.\end{aligned}$$

For $a \geq 2$,

$$\begin{aligned}E_a &\ll \binom{j}{a} \left\| \frac{X}{\sqrt{n}} \right\|^{j-a} \left\| \frac{X_L}{\sqrt{n}} \right\|^{a-1} \|X_L\| \\ &= o(1) \text{ w.h.p.,}\end{aligned}$$

where we have used (20) and that $a \geq 2$.

To bound E_1 , we have

$$\begin{aligned}E_1 &\leq \sum_{m=0}^{j-1} \left| u_i^* \left(\frac{X}{\sqrt{n}} \right)^m X_L \left(\frac{X}{\sqrt{n}} \right)^{j-1-m} v_i \right| \\ &\leq \sum_{(k,l) \in L} \sum_{m=0}^{j-1} |x_{kl}| \left| u_i^* \left(\frac{X}{\sqrt{n}} \right)^m e_k \right| \left| e_l^T \left(\frac{X}{\sqrt{n}} \right)^{j-1-m} v_i \right| \\ &\ll \frac{\delta_n}{\sqrt{n}} \sum_{(k,l) \in L} |x_{kl}| \text{ w.h.p.}\end{aligned}$$

Note that if $j \geq 2$, then either $m \geq 1$ or $j-1-m \geq 1$ for $0 \leq m \leq k-1$. Hence the last line follows from (19).

Since $\mathbb{E}|x_{kl}| \leq 1$, $\delta_n = \log n$ and $|L| = O(n^{1/2-2\delta})$, we have $E_1 \rightarrow_P 0$ by Markov's inequality, and (17) follows.

2.1.2. *Step 2.* We first state and prove the complex version of Wick's theorem (also known as Isserlis' theorem, see [10]) which will be needed later.

Lemma 4. (Complex Wick's theorem)

Let $(Z_1, Z_2, \dots, Z_n) = (X_1 + iY_1, \dots, X_n + iY_n)$ be a centered complex Gaussian vector. Thus the vector $(X_1, Y_1, \dots, X_n, Y_n)$ is multivariate normal. Then for any $I = (i_1, \dots, i_{2k}) \in [n]^{2k}$,

$$E \prod_{l=1}^{2k} Z_{i_l} = \sum_P \prod_{j=1}^k \mathbb{E}[Z_{i_{p_{2j-1}}} Z_{i_{p_{2j}}}]$$

where the sum is over all partitions $P = \bigcup_{j=1}^k \{p_{2j-1}, p_{2j}\}$ of $[2k]$ into pairs. Also, the left hand side is 0 if I has odd length.

Proof. Wick's theorem is the statement of the lemma for multivariate centered real Gaussians. The complex version follows by expanding both sides of the equation into real and imaginary parts and applying Wick's theorem. Let

$$W_i^a = \begin{cases} X_i & : a = 1 \\ iY_i & : a = 2 \end{cases}$$

Then

$$\mathbb{E} \prod_{l=1}^{2k} Z_{i_l} = \sum_{a_1, \dots, a_{2k} \in \{1, 2\}} \prod_{l=1}^{2k} W_{i_l}^{a_l}$$

while

$$\sum_P \prod_{i=1}^k \mathbb{E}[Z_{p_{2i-1}} Z_{p_{2i}}] = \sum_P \sum_{a_1, \dots, a_{2k} \in \{1, 2\}} \prod_{i=1}^k \mathbb{E}[W_{p_{2i-1}}^{a_{2i-1}} W_{p_{2i}}^{a_{2i}}].$$

Switching the sums and applying Wick's theorem to $\mathbb{E} \prod_{l=1}^{2k} W_{i_l}^{a_l}$ for each choice of the a_l 's yields the result. □

We now prove Proposition 1 for the collection of random variables $(Z_{i,1}^{L^c})_{i=1}^p \cup (Z_{i,j})_{i=1,j=2}^{p,m}$. This part of the proof employs the moment method in a similar way to those in [20] and [6]. To avoid notational clutter on a first reading, one may set $p = 1$ to grasp the main ideas of the proof.

To handle the $j = 1$ case uniformly, in the proof we will abuse notation by writing $Z_{i,1}$ for $Z_{i,1}^{L^c}$ and $G_{i,1}$ for $G_{i,1}^{L^c}$. When $j = 1$, we will denote X_{L^c} by X^j and finally, we define

(21)

$$C_{i_1, i_2}^{(d_1), (d_2)}(j) := \begin{cases} C_{i_1, i_2}^{(d_1), (d_2)} : j \geq 2 \\ \lim_{n \rightarrow \infty} \sum_{(k, l) \in L^c} \overline{(u_{i_1, k})}^{(d_1)} \overline{(u_{i_2, l})}^{(d_2)} (v_{i_1, l})^{(d_1)} (v_{i_2, l})^{(d_2)}. \end{cases}$$

By Carleman's theorem for the case of a complex vector of random variables (see e.g. [1]), it suffices to show that the multivariate mixed moments converge. Namely,

$$(22) \quad \mathbb{E} \prod_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}} Z_{i,j}^{r_{i,j}} \overline{Z_{i,j}}^{s_{i,j}} = \mathbb{E} \prod_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}} G_{i,j}^{r_{i,j}} \overline{G_{i,j}}^{s_{i,j}} + o(1)$$

for $(r_{i,j})_{i=1,j=1}^{p,m}, (s_{i,j})_{i=1,j=1}^{p,m} \in \mathbb{N}^{pm}$.

Let $Q_1 := -\frac{1}{2} \sum_{i,j} (j-1)(r_{i,j} + s_{i,j})$. Then the left hand side of (22) is

$$(23) \quad n^{-Q_1} \mathbb{E} \prod_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}} (u_i^* X^j v_i)^{r_{i,j}} (u_i^T \overline{X^j v_i})^{s_{i,j}}.$$

Expanding the product in (23) will yield terms corresponding to the union of directed paths on the vertex set $[n]$ with $\sum_i r_{i,j} + s_{i,j}$ of them having length j for each $1 \leq j \leq m$. We first introduce notation in order to write (23) as a sum $n^{-Q_1} \sum_* W(F)$, with $*$ and $W(F)$ defined appropriately. Next, we reduce the sum to terms with paths having multiplicity two and disjoint interior vertices (see Lemma 5). Finally we apply the complex Wick theorem to obtain the proposition.

Let

$$S := \{(a, b, c, d) : a \in [p], b \in [m], d \in \{0, 1\}, c \in [r_{a,b}] \text{ if } d = 0 \text{ and } c \in [s_{a,b}] \text{ if } d = 1\}$$

be the index set for the $Z_{i,j}$'s. For $s \in S$ we write $s = (s_a, s_b, s_c, s_d)$. Recalling (1), (23) can be written as

$$(24) \quad n^{-Q_1} \mathbb{E} \prod_{s \in S} (u_{s_a}^* X^{s_b} v_{s_a})^{(s_d)}.$$

We let

$$T := \{(s, e) : s \in S \text{ and } e \in [s_{b+1}]\}$$

be the index set of terms within the $Z_{i,j}$'s. For $t \in T$, we write

$$t = (t_s, t_e) = (t_a, t_b, t_c, t_d, t_e).$$

By a slight abuse of notation, we will write u_t for u_{t_a} and u_s for u_{s_a} . We denote the index set for terms in the expansion of (24) by

$$\mathcal{F}' := \{F : T \rightarrow [n] : t_b = 1 \Rightarrow (F(t, 1), F(t, 2)) \in L^c\}.$$

Finally for $s \in S$ and $F \in \mathcal{F}'$ let

$$(25)$$

$$W_s(F) := (u_{s,F(s,1)}^* v_{s,F(s,s_b+1)} \mathbb{1}_{[s_b \geq 2 \text{ or } (F(s,1), F(s,2)) \in L^c]})^{(s_d)} (\mathbb{E} \prod_{e=1}^{s_b} x_{F(s,e), F(s,e+1)})^{(s_d)}$$

$$(26)$$

$$=: W_{s,(u,v)}(F) W_{s,x}(F)$$

and set

$$(27) \quad W_{u,v}(F) := \prod_{s \in S} W_{s,(u,v)}(F),$$

$$W_x(F) := \prod_{s \in S} W_{s,x}(F)$$

and

$$(28) \quad W(F) := \prod_{s \in S} W_s(F).$$

Now we can write (24) as

$$(29) \quad n^{-Q_1} \mathbb{E} \prod_{s \in S} (u_s^* X^{s_b} v_s)^{(s_d)} = n^{-Q_1} \sum_{F \in \mathcal{F}'} W(F).$$

For each partition $\mathcal{T} = \{T_1, \dots, T_q\}$ of T , set

$$\mathcal{F}_{\mathcal{T}} := \{F \in \mathcal{F} : \{F^{-1}(i) : i \in [n], F^{-1}(i) \neq \emptyset\} = \{T_1, \dots, T_q\}\}$$

to be the set of terms F whose preimages induce the partition $\{T_1, \dots, T_q\}$. We can now write

$$\sum_{F \in \mathcal{F}} W(F) = \sum_{\mathcal{T} = \{T_1, \dots, T_q\}} \sum_{F \in \mathcal{F}_{\mathcal{T}}} W(F).$$

We now define notation for the edges of the graph induced by the terms F . First, let $E := \{(t, t') \in T^2 : t_s = t'_s, t'_e = t_e + 1\}$ and fix a partition $\mathcal{T} = \{T_1, \dots, T_q\}$ of T . For $F \in \mathcal{F}_{\mathcal{T}}$ and $i, j \in [q] = [q(\mathcal{T})]$, let

$$E_{i,j}^{\mathcal{T}} := \{e = (t, t') \in E : t \in T_i \text{ and } t' \in T_j\}$$

and let

$$E_{\mathcal{T}} := \{E_{i,j}^{\mathcal{T}} : |E_{i,j}^{\mathcal{T}}| > 0\}.$$

Note that $(|e|)_{e \in E_{\mathcal{T}}}$ is independent of $F \in \mathcal{F}_{\mathcal{T}}$ and that

$$(30) \quad W_x(F) = \prod_{e \in E_{\mathcal{T}}} \mathbb{E}|x|^{|e|}.$$

Since $\mathbb{E}|x| = 0$, $W_x(F) = 0$ if $|e| = 0$ for any $e \in E_{\mathcal{T}}$. Thus defining

$$\mathcal{F} := \bigcup_{\substack{\mathcal{T} \text{ partition of } T: \\ |e| \geq 2 \forall e \in E_{\mathcal{T}}}} \mathcal{F}_{\mathcal{T}},$$

we have

$$(31) \quad n^{-Q_1} \sum_{F \in \mathcal{F}'} W(F) = n^{-Q_1} \sum_{F \in \mathcal{F}} W(F).$$

Each $F \in \mathcal{F}$ can be interpreted as a union of paths on $[n]$. More precisely, letting $T_s := \{t \in T : t_s = s\}$, we define $\pi_{F,s} := F|_{T_s}$ to be the path of F corresponding to term $s \in S$. The interior vertices of $\pi_{F,s}$ are defined to be $F(\{(s, e) : e = 2, 3, \dots, s_b\})$.

Lemma 5. *Assume the hypotheses of Proposition 1 and recall the notation introduced above. Let \mathcal{F}_0 be the set of terms F such that each path $\pi_{F,s}$ for $s \in S$ has multiplicity 2 and different paths have disjoint interior vertices.*

Then

$$n^{-Q_1} \sum_{F \in \mathcal{F}} W(F) = n^{-Q_1} \sum_{F \in \mathcal{F}_0} W(F) + o(1).$$

We will postpone the proof of the lemma to the end of the section. Assuming the lemma, we now prove the proposition.

First suppose $\sum_i r_{i,j} + s_{i,j}$ is odd for some j . Then \mathcal{F}_0 is empty and the left-hand side of (22) is $o(1)$ which matches the right-hand side by the vanishing of odd mixed moments of a centered complex Gaussian. For the rest of the proof, we can thus assume that for each j , $\sum_i r_{i,j} + s_{i,j}$ is even.

We group the terms in \mathcal{F}_0 as follows. Let $S_j := \{s \in S : s_b = j\}$ and define \mathcal{P}_j to be the set of unordered partitions of S_j into parts of size two. Note that by assumption, $|S_j|$ is even for all j .

For $F \in \mathcal{F}_0$, note by (25) and (28) that $W(F)$ does not depend on the interior points $\{f(s, e) : s \in S, e = 2, \dots, s_b\}$. There are $\sum_{i,j} (j-1)(r_{i,j} + s_{i,j})$ such points which occur in pairs and can be chosen in n^{Q_1} ways.

For $F \in \mathcal{F}_0$ and $j \in [m]$, let $P_{F,j} \in \mathcal{P}_j$ be the partition of \mathcal{P}_j induced by F . Then F satisfies the condition that for each part $\{p, q\} \in P_{F,j}$, $F(p, 1) = F(q, 1)$ and $F(p, p_b + 1) = F(q, q_b + 1)$.

Summing over the choices for interior points and F satisfying the above condition instead of summing over $F \in \mathcal{F}_0$ incurs an $o(1)$ error and we have

$$(32) \quad n^{-Q_1} \sum_{F \in \mathcal{F}_0} W(F) = \prod_{j=1}^m \sum_{P_j \in \mathcal{P}_j} \prod_{\{p,q\} \in P_j} \sum_{\substack{F(p,1)=F(q,1), \\ F(p,j+1)=F(q,j+1) \in [n]}} W_p(F) W_q(F) + o(1)$$

where, recalling (25),

$$W_p(F) W_q(F) = (\mathbb{E} x^{(p_d)} x^{(q_d)})^j \prod_{r \in \{p,q\}} \bar{u}_{r,F(r,1)}^{(r_d)} v_{r,F(r,j+1)}^{(r_d)} \mathbb{1}_{[j \geq 2 \text{ or } (F(r,1), F(r,2)) \in L^c]}.$$

Finally, using (13) and (21), (32) evaluates to

$$(33) \quad \prod_{j=1}^m \sum_{P_j \in \mathcal{P}_j} \prod_{\{p,q\} \in P_j} (\mathbb{E} x^{(p_d)} x^{(q_d)})^j C_{p_a, q_a}^{(p_d), (q_d)}(j) + o(1).$$

On the other hand, we let \mathcal{P} be the set of partitions of S into pairs and for $s \in S$, we set $G_s := G_{s_a, s_b}^{(s_d)}$. Note that for $j \neq k$, $\mathbb{E} G_{i_1, j} G_{i_2, k} = 0$ and hence $G_{i_1, j}$ and $G_{i_2, k}$

are independent. Applying Wick's theorem to the right hand side of (22) gives

$$\begin{aligned} \mathbb{E} \prod_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}} G_{i,j}^{r_{i,j}} \overline{G_{i,j}^{s_{i,j}}} &= \mathbb{E} \prod_{s \in S} G_s \\ &= \prod_{j=1}^m \mathbb{E} \prod_{s \in S_j} G_s \\ &= \prod_{j=1}^m \sum_{P_j \in \mathcal{P}_j} \prod_{\{p,q\} \in P_j} \mathbb{E} G_p G_q \end{aligned}$$

where we have used Wick's theorem in the third line. Comparing (33) and (15) then concludes the proof of the proposition.

Note the following special cases of Proposition 1, where we write $G_{i,1}$ for $G_{i,1}^{L^c}$.

(i) If $\mathbb{E}x^2 = 0$, condition (15) becomes

$$(34) \quad \mathbb{E} G_{i_1,j} \overline{G_{i_2,k}} = \delta_{jk} C^{(0),(1)}(i_1, i_2)$$

and

$$\mathbb{E} G_{i_1,j} G_{i_2,k} = 0.$$

(ii) If we further assume that for $p = d^2$, the vectors $(u_i, v_i)_{i=1}^p$ are of the form $(u_a, u_b)_{a,b=1}^d$ with $(u_a)_{a=1}^d$ orthonormal, then (34) reduces to

$$(35) \quad \mathbb{E} G_{(a,b),j} \overline{G_{(c,d),k}} = \delta_{jk} \delta_{ab} \delta_{cd}.$$

2.2. Proof of Lemma 5. Fix a partition $\mathcal{T} = \{T_1, \dots, T_q\}$ of T with $|e| \geq 2$ for every $e \in E_{\mathcal{T}}$. We first rewrite the sum $n^{-Q_1} \sum_{F \in \mathcal{F}_{\mathcal{T}}} W(F)$ as a product of terms over $j \in [q]$.

Define $T^1 := \{t \in T : t_e = 1\}$, $T^2 := \{t \in T : t_e = t_b + 1\}$, $T^3 := T \setminus (T^1 \cup T^2)$ and let $T_j^l := T_j \cap T^l$ for $l = 1, 2, 3$. For $t \in T$ and $i \in [n]$, define the vertex weights

$$(36) \quad w(t, i) := \begin{cases} |u_{t,i}| & : t \in T^1 \\ |v_{t,i}| & : t \in T^2 \\ n^{-1/2} & : t \in T^3 \end{cases} \quad ..$$

The $w(t, i)$'s account for the factors n^{-Q_1} and $W_{u,v}(F)$ in (27) and (29) respectively. Since $\mathbb{E}|x|^a \ll K^{(a-4)_+}$, using (30) we have

$$\begin{aligned} n^{-Q_1} \left| \sum_{F \in \mathcal{F}_{\{T_1, \dots, T_q\}}} W(F) \right| &\leq \sum_{\substack{i_1, \dots, i_q \in [n] \\ \text{distinct}}} \left(\prod_{j=1}^q \prod_{t \in T_j} w(t, i_j) \right) \prod_{e \in E_{\mathcal{T}}} K^{(|e|-4)_+} \\ (37) \quad &\leq \sum_{i_1, \dots, i_q \in [n]} \left(\prod_{j=1}^q \prod_{t \in T_j} w(t, i_j) \right) \prod_{e \in E_{\mathcal{T}}} K^{(|e|-4)_+}. \end{aligned}$$

We would like to bound $\prod_{e \in E_{\mathcal{T}}} K^{(|e|-4)+}$ by $\prod_{t \in T} K^*(t, i_{j(t)})$ for some suitably defined K^* in order to bound the right-hand side (37) by

$$\prod_{j=1}^q \sum_{i_1, \dots, i_q \in [n]} \prod_{t \in T_j} w(t, i_j) K^*(t, i_j).$$

We do this first for the expression $\prod_{e \in E_{\mathcal{T}}} K^{|e|}$ in order to motivate some of the technical definitions. Fix $i_1, \dots, i_q \in [n]$ and assume for $t \in T$ and $j \in [q]$ that $|u_{t, i_j}|, |v_{t, i_j}| \neq 0$. Recall the parameter $c \in [0, 1]$ from (2.1). For $t \in T_j$, $t^1 \in T_j^1$ and $t^2 \in T_j^2$, define

$$(38) \quad K(t, i) := \begin{cases} |u_{t, i}^{-(1-\epsilon)}| & : t \in T^1 \\ |v_{t, i}^{-(1-\epsilon)}| & : t \in T^2 \\ \max(K^2 n^{-c(1-\epsilon)}, K) & : t \in T^3 \end{cases}.$$

We first show that

$$(39) \quad \prod_{e \in E_{\mathcal{T}}} K^{|e|} \ll \prod_{j=1}^q \prod_{t \in T_j} K(t, i_j).$$

Fix $s \in S$. Suppose $s_b = 1$. Then for δ and ϵ sufficiently small,

$$(40) \quad \begin{aligned} \prod_{t \in T: t_s = s} K(t, i_{j(t)}) &\geq \min_{(k, l) \in L^c} |u_{t, k} v_{t, l}|^{-(1-\epsilon)} \\ &\gg n^{\max(1/4 - \delta, c)(1-\epsilon)} \\ &\gg K. \end{aligned}$$

The last line follows from $M < \max(1/4, c)$ which is a consequence of (16), .

If $s_b \geq 2$,

$$(41) \quad \begin{aligned} \prod_{t \in T: t_s = s} K(t, i_{j(t)}) &\gg K^2 n^{c(1-\epsilon)} (\|u_t\|_{\infty} \|v_t\|_{\infty})^{-(1-\epsilon)} K^{s_b-2} \\ &\gg K^{s_b} \end{aligned}$$

where we have used $\|u_t\|_{\infty} \|v_t\|_{\infty} \ll n^{-c}$. Using (40) and (41) and taking the product over $s \in S$ gives (39). We now define $K^*(t, i)$ in such a way that we have the analogous bound

$$(42) \quad \prod_{e \in E_{\mathcal{T}}} K^{(|e|-4)+} \ll \prod_{j=1}^q \prod_{t \in T_j} K^*(t, i_j).$$

First, order the elements of $T_j^l = \{t_1^l, t_2^l, \dots, t_{|T_j^l|}^l\}$ arbitrarily for $l = 1, 2, 3$. We define the set $C_j \subset T_j$ by the following conditions.

- (i) $t_k^3 \in C_j \iff k \leq 2$.
- (ii) For $l = 1, 2$, $t_k^l \in C_j \iff k + |T_j^3| \leq 2$.

It is easy to verify that $|C_j \setminus T_j^1|, |C_j \setminus T_j^2| \leq 2$. We now define

$$K^*(t, i) := \begin{cases} 1 : t \in C_j \\ K(t, i) : \text{otherwise} \end{cases}.$$

We now prove (42). Fix $e \in E_{\mathcal{T}}$ and suppose $e \subset T_i \times T_j$. Define $e' \subset e$ by

$$e' := \{(s, t) \in e : s \in C_i \text{ or } t \in C_j\}.$$

Since $|C_i \setminus T_i^2|, |C_j \setminus T_j^1| \leq 2$, $|e'| \leq 4$ and we have

$$\prod_{e \in \mathcal{T}} K^{|e|-4} \leq \prod_{e \in \mathcal{T}} K^{|e \setminus e'|}.$$

It thus suffices to show

$$\prod_{e \in \mathcal{T}} K^{|e \setminus e'|} \leq \prod_{j=1}^q \prod_{t \in T_j} K^*(t, i_j).$$

As in the proof of (39), we fix $s \in S$. Let $C := \bigcup_{j \in [q]} C_j$ and define

$$e_s = \{(s, l), (s, l+1) : 1 \leq l \leq s_b \text{ and } (s, l), (s, l+1) \notin C\}$$

and

$$v_s = \{(s, l) : (s, l) \notin C \text{ and } l = 2, 3, \dots, s_b\}.$$

Since $K(t, i) \geq 1$ for $t = (s, 1)$ and $t = (s, s_b + 1)$, it suffices to show

$$K^{|e_s|} \leq \prod_{t \in v_s} K(t, i).$$

If $|e_s| = s_b$, this follows from (40) and (41). Now suppose $|e_s| < s_b$. We first show that $|e_s| \leq |v_s|$. Choose l^* such that $(s, l^*) \in C$ and define the map $f : e_s \rightarrow v_s$ by

$$f((s, l), (s, l+1)) := \begin{cases} (s, l+1) : l \leq l^* - 2 \\ (s, l) : l \geq l^* + 1 \end{cases}.$$

We see that f is injective and hence $|e_s| \leq |v_s|$. Since $K(t, i) \geq K$ for $t \in v_s$, we have

$$\begin{aligned} K^{|e_s|} &\leq K^{|v_s|} \\ &\leq \prod_{t \in v_s} K(t, i) \end{aligned}$$

completing the proof of (42).

We can now use (42) in (37) to write

$$\begin{aligned} \left| n^{-Q_1} \sum_{F \in \mathcal{F}_{\{T_1, \dots, T_q\}}} W(F) \right| &\leq \sum_{i_1, \dots, i_q \in [n]} \prod_{j=1}^q \prod_{t \in T_j} w(t, i_j) K^*(t, i_j) \\ (43) \quad &= \prod_{j=1}^q \left(\sum_{i_j \in [n]} \prod_{t \in T_j} w(t, i_j) K^*(t, i_j) \right) \end{aligned}$$

$$(44) \quad =: \prod_{j=1}^q W^*(T_j).$$

We now fix a part of \mathcal{T} , say T_1 and consider $W^*(T_1)$. To prove Lemma 5, it suffices to prove the following.

- Lemma 6.** (i) $W^*(T_1) = O(1)$
(ii) If $|T_1^3| \geq 1$, then $|W^*(T_1)| = o(1)$ unless $|T_1^3| = |T_1| = 2$.
(iii) $\prod_j W^*(T_j) = o(1)$ unless $|e| = 2$ for every $e \in E_{\mathcal{T}}$.

Proof. We first show that

$$(45) \quad w(t, i)K(t, i) = \begin{cases} O(1) : t \in T_1^1 \cup T_1^2 \\ o(1) : t \in T_1^3 \end{cases}$$

using (36), (38) and (16). Suppose $t \in T_1^1$. Then $w(t, i)K(t, i) \leq |u_i|^\epsilon = O(1)$. We have a similar bound for $t \in T_1^2$. Finally, if $t \in T_1^3$, then

$$w(t, i)K(t, i) = n^{-1/2} \max(K^2 n^{-c(1-\epsilon)}, K).$$

Since $K = o(n^M)$ and $M \leq \min(1/2, c)$, we have the desired bound. This implies in particular that for any $D \subset C_1$,

$$(46) \quad W^*(T_1) \ll \sum_{i \in [n]} \prod_{t \in D} w(t, i).$$

We prove Lemma 6.(ii) first. For u and v unit vectors in \mathbb{C}^n , we will need the estimate

$$(47) \quad \sum_{i \in [n]} |u_i|^\epsilon \ll O(n^{1-\epsilon/2})$$

which follows from Hölder's inequality. Suppose $|T_1^3| = 1$. Then, since each edge has multiplicity at least 2, we must have $|T_1^1|, |T_1^2| \geq 1$. Applying (46) with $D = C_1 = \{t_1^1, t_1^2, t_1^3\}$, we have that for some (u, v) ,

$$\begin{aligned} W^*(T_1) &\ll \sum_{i=1}^n n^{-1/2} |u_i| |v_i| \\ &\leq O(n^{-1/2}). \end{aligned}$$

If $|T_1^3| \geq 3$, then $C_1 = \{t_1^3, t_2^3\}$ and

$$\begin{aligned} W^*(T_1) &\ll \sum_{i \in [n]} n^{-1} K^*(t_3^3, i) \\ &= o(1) \end{aligned}$$

by (45). Finally, suppose $|T_1^3| = 2$. Suppose $|T_1^1| \geq 1$. Then from (47), we have

$$W^*(T_1) \ll \sum_{i \in [n]} n^{-1} |u_i|^\epsilon = o(1).$$

We have a similar estimate if $|T_1^2| \geq 1$. We conclude that if $|T_1^3| \geq 1$, $W^*(T_1) = o(1)$ unless $|T_1^3| = 2$ and $|T_1^1| = |T_1^2| = 0$, in which case $W^*(T_1) = O(1)$.

We now prove (iii). Assume first that e is an edge incident to distinct vertices, say $e \subset T_1 \times T_2$, and that $|e| \geq 3$. By (ii), we may assume $T_1^3 = T_2^3 = \emptyset$. Since $|T_1^1|, |T_2^2| \geq 3$, we may choose $(s_i, t_i) \in e$ for $i = 1, 2, 3$ where $s_i \in T_1^1$ and $t_i \in T_2^2$

and let $C_1 = \{s_1, s_2\}$ and $C_2 = \{t_1, t_2\}$. Then bounding $W^*(T_1)W^*(T_2)$ by the contribution from $(s_i, t_i)_{i=1}^3$, we have

$$\begin{aligned} W^*(T_1)W^*(T_2) &\ll \sum_{(i,j) \in L^c} \prod_{k=1}^2 |u_{s_k, i} v_{t_k, j}| |u_{s_3, i} v_{t_3, j}|^\epsilon \\ &\leq \max_{(i,j) \in L^c} |u_{s_3, i} v_{t_3, j}|^\epsilon \sum_{i \in [n]} |u_{s_1, i}| |u_{s_2, i}| \sum_{j \in [n]} |v_{t_1, j}| |v_{t_2, j}| \\ &= o(1). \end{aligned}$$

We have a similar bound if e is a loop at say T_1 .

To complete the proof of the lemma, it remains to prove (i) in the cases not covered by (ii) and (iii). Thus, set $|T_1^3| = 0$ and assume without loss of generality that $|T_1^1| \geq 2$. Then with $D = \{t_1^1, t_1^2\} =: \{s, t\}$ in (46) we have

$$\begin{aligned} W^*(T_1) &\ll \sum_{i \in [n]} |u_{s, i} u_{t, i}| \\ &= O(1). \end{aligned}$$

□

3. PROOF OF LEMMA 7

Recall the bilinear average of the normalized resolvent introduced in (12) in Section 2. In this section we control the tail of its Neumann series and, with the help of Proposition 1, obtain the joint limiting distribution of such terms in Lemma 7. This is the main ingredient in the proof of Theorem 4 which is presented in the next section.

Lemma 7. *Fix complex numbers $\theta_1, \dots, \theta_a$ with $|\theta_j| > 1$ for $j \in [a]$ and suppose $\lambda_j = \lambda_{n,j} \rightarrow_P \theta_j$ as $n \rightarrow \infty$. Let $(u_i, v_i)_{i=1}^p$ be p pairs of vectors satisfying the hypotheses of Proposition 1. Let*

$$S_{i,j} := \sum_{k \geq 1} \frac{\sqrt{n} \left\langle \left(\frac{X}{\sqrt{n}} \right)^k v_i, u_i \right\rangle}{\lambda_j^k} =: \sum_{k \geq 1} \frac{Z_{i,k}}{\lambda_j^k}.$$

Recall the definition of $(G_{i,1})_{i=1}^p$ from Proposition 1 and define centered complex Gaussians $(g_{i,j})_{i=1,j=1}^{p,a}$ independent of $(G_{i,1})_{i=1}^p$ with mixed second moments given by

$$(48) \quad \mathbb{E} g_{i,j}^{(d_1)} g_{i',j'}^{(d_1)} = \frac{(\mathbb{E} x^{(d_1)} x^{(d_2)})^2}{\theta_j \theta_{j'} (\theta_j \theta_{j'} - \mathbb{E} x^{(d_1)} x^{(d_2)})} U_{i,i'}^{(d_1), (d_2)} V_{i,i'}^{(d_1), (d_2)}.$$

Then

$$(S_{i,j})_{i=1,j=1}^{p,a} \Rightarrow (F_{i,j})_{i=1,j=1}^{p,a}$$

where

$$(49) \quad F_{i,j} := \frac{G_{i,1}}{\theta_j} + g_{i,j}.$$

To prove the lemma, we split $S_{i,j}$ into three sums as follows. Fix cutoffs $m > 0$ and $T_n = \log^2 n$ ($T_n = \omega(\log n)$ suffices) and define

$$\begin{aligned} S_{i,j} &= \sum_{k=1}^m \frac{Z_{i,k}}{\lambda_j^k} + \sum_{k=m+1}^{T_n} \frac{Z_{i,k}}{\lambda_j^k} + \sum_{k>T_n}^{\infty} \frac{Z_{i,k}}{\lambda_j^k} \\ &=: S_{i,j}^A + S_{i,j}^B + S_{i,j}^C. \end{aligned}$$

We define

$$(50) \quad T_{i,j}^A := \sum_{k=1}^m \frac{G_{i,k}}{\theta_j^k}$$

where the $G_{i,k}$ are defined as in the statement of Proposition 1. Note that $T_{i,j}^A$ is independent of n .

By Proposition 1 and the multivariate version of Slutsky's theorem (see [5]),

$$((Z_{i,k}), (\lambda_j)) \Rightarrow ((G_{i,k}), (\theta_j)),$$

where the joint convergence is over all $i \in [p]$, $k \in [m]$ and $j \in [a]$. By the continuous mapping theorem, $(S_{i,j}^A) \Rightarrow (T_{i,j}^A)$ jointly for $i \in [p]$ and $j \in [a]$. By the definitions of $T_{i,j}^A$ in (50) and of $G_{i,k}$ in (14) and (15), and by inspecting (48) and (49), we see that

$$T_{i,j}^A \xrightarrow{m \rightarrow \infty} F_{i,j}$$

jointly.

To prove Lemma 7, it suffices to prove

Lemma 8. (a) $\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}|S^B| = 0$ and
(b) $\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}|S^C| = 0$.

where we have suppressed the i and j dependence for $S_{i,j}^B$ and $S_{i,j}^C$.

Define the event

$$(51) \quad E_n := \{|\lambda_{n,j} - \theta_j| < \delta_j := \frac{|\theta_j| - 1}{4} \text{ for all } j \in [a]\}.$$

By hypothesis $\mathbb{P}(E_n) = 1 - o(1)$ so it suffices to prove Lemma 7 (and hence Lemma 8) on E_n . In the following, we fix an index j and set $\delta := \frac{|\theta| - 1}{4}$. Note that we have

$$(52) \quad |\lambda| > 1 + \frac{3}{4}(|\theta| - 1).$$

We prove Lemma 8b first.

Proof. Recall that on E_n , $|\lambda| > 1 + 3\delta$ (see (51)). By Theorem 2, $\rho(X/\sqrt{n}) < 1 + \delta$ w.h.p. and we can choose l such that $\|(\frac{X}{\sqrt{n}})^l\|^{1/l} < 1 + 2\delta$. We may assume without loss of generality that these events occur on E_n . By submultiplicativity of the

operator norm,

$$\begin{aligned} \left\| \left(\frac{X}{\sqrt{n}} \right)^k \right\| &\leq \left\| \left(\frac{X}{\sqrt{n}} \right)^l \right\|^{\lfloor \frac{k}{l} \rfloor} \max_{0 \leq i < l} \left\| \left(\frac{X}{\sqrt{n}} \right)^i \right\| \\ &\leq O_l(1 + 2\delta)^k \text{ w.h.p.} \end{aligned}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |S_C| &\leq \sum_{k > T_n} \frac{\sqrt{n} \left\| \left(\frac{1}{\sqrt{n}} X \right)^k \right\| |u|_2 |v|_2}{|\lambda|^k} \\ &< O_l(\sqrt{n}) \sum_{k > T_n} \left(\frac{1 + 2\delta}{1 + 3\delta} \right)^k \\ &= o(1) \end{aligned}$$

where the last line follows from our choice of $T_n = \log^2 n$. \square

To prove Lemma 8a, we will need

Lemma 9. *Let u and v be unit vectors in \mathbb{C}^n and set*

$$Z_k := \sqrt{n} u^* \left(\frac{1}{\sqrt{n}} X \right)^k v.$$

Fix $\epsilon > 0$ and assume $|x| \leq K = O(n^{\frac{1-\epsilon}{2}})$. Then there exists $c = c(\epsilon) > 0$ such that for all $k \ll n^c$,

$$(53) \quad \mathbb{E}|Z_k|^2 = O(1).$$

Assuming Lemma 9 we prove Lemma 8a on E_n . Since $(\mathbb{E}|Z|)^2 \leq \mathbb{E}|Z|^2$, we have

$$\begin{aligned} \mathbb{E}|S^B| &\leq \sum_{k=m+1}^{T_n} \mathbb{E} \frac{|Z_k|}{|\lambda|^k} \\ &\ll \left| 1 + \frac{3}{4}(|\theta| - 1) \right|^{-2m} \end{aligned}$$

where we have used Lemma 9 and (52) in the last line. Lemma 8(a) follows from letting $m \rightarrow \infty$.

Remark 6. *Note that by the truncation argument given in Appendix B, Lemma 9, and hence Lemma 8a, is valid under the moment hypothesis $\mathbb{E}|x|^{4+\epsilon} < \infty$ for any fixed $\epsilon > 0$.*

3.1. Proof of Lemma 9. In this subsection we prove Lemma 9.

Proof. It suffices to show

$$(54) \quad \mathbb{E}|u^* X^k v|^2 = O(n^{k-1}).$$

Let

$$T := \{(a, b) : a = 1, 2, b = 0, 1, \dots, k\},$$

$$T' := \{(a, b) \in T : b < k\}$$

and

$$E := \{((a, b), (a, b + 1)) \in T^2 : b < k\}.$$

Let $T_P := T|_{a=1}$, $T_Q := T|_{a=2}$ and for $t \in T'$, set $t^s := (a, b + 1)$. We will designate the terms in the expansion of (54) by

$$\mathcal{P}' := \{F : T \rightarrow [n]\}.$$

For $F \in \mathcal{P}'$, let $F_P := F|_{T_P}$ and $F_Q := F|_{T_Q}$. Let

$$W_{u,v}(F) := |u_{F(1,0)} u_{F(2,0)} v_{F(1,k)} v_{F(2,k)}|$$

and

$$W_x(F) := \mathbb{E} \left| \prod_{t \in T'} x_{F(t), F(t^s)} \right|.$$

Then we have

$$(55) \quad \mathbb{E} |u^* X^k v|^2 \leq \sum_{F \in \mathcal{P}'} W_{u,v}(F) W_x(F).$$

For $F \in \mathcal{P}'$, let

$$E_F := \{(F(t), F(t^s)) \in [n]^2 : t \in T'\}.$$

denote the edges of F and let

$$\mathcal{E}_F := \{t \in T' : (F(t), F(t^s)) = (i, j) : (i, j) \in E^F\}.$$

Then

$$W_x(F) = \prod_{e \in \mathcal{E}_F} \mathbb{E} |x|^{|e|}.$$

Noting that $\mathbb{E} |x| = 0$ and letting

$$\mathcal{P} := \{F \in \mathcal{P}' : |e| \geq 2 \text{ for all } e \in \mathcal{E}_F\},$$

we have

$$(56) \quad \mathbb{E} |u^* X^k v|^2 \leq \sum_{F \in \mathcal{P}} W_{u,v}(F) W_x(F).$$

Now, for a fixed $F \in \mathcal{P}$, let

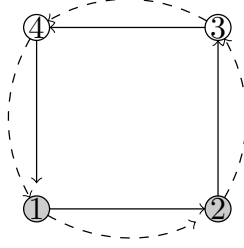
$$V = V_F := \{F(t) : t \in T\}$$

be the set of vertices. For $v \in V$ let $m(v) = |F^{-1}(v)|$ denote its multiplicity. Let $d_{\text{in}}(v) := |\{x \in [n] : (x, v) \in E\}|$ and $d_{\text{out}}(v) := |\{x : (v, x) \in E\}|$ denote its indegree and outdegree. Finally, let $d(v) := d_{\text{in}}(v) + d_{\text{out}}(v)$ be the (total) degree of v .

Shown in Figure 3 is an example with $k = 4$ with the paths $(1, 2, 3, 4)$ and $(2, 3, 4, 1)$. Each vertex has indegree 2 and outdegree 2.

We will first determine the main term from \mathcal{P} and its contribution to (56).

Lemma 10. *Suppose $F \in \mathcal{P}$. Then $|V| \leq k + 1$ and that equality occurs only when $F_P = F_Q$ and $|F_P| = |F_Q| = k + 1$.*

FIGURE 3. An example of $F \in \mathcal{P}$ with $k = 4$.

Fix $v \in V_F$ and suppose $d(v) = 1$. Since each edge has multiplicity at least two, we have the following.

- (i) If $d_{\text{in}}(v) = 1$, $v = F(1, 0) = F(2, 0)$.
- (ii) If $d_{\text{out}}(v) = 1$, $v = F(1, k) = F(2, k)$

In particular, if two vertices of V have degree 1, then one has outdegree 1, the other has indegree 1 and the rest have both outdegree and indegree of at least 1. Since $|e| \geq 2$ for each $e \in \mathcal{E}_F$, we also have $|E_F| \leq k$. Thus

$$\begin{aligned} 2k \geq 2|E_F| &= \sum_{v \in V} d(v) \\ &\geq 1 + 1 + 2(|V| - 2) \\ &= 2(|V| - 1). \end{aligned}$$

Thus, $|V| \leq k + 1$ with equality occurring only when two of the vertices have degree 1 and the rest have degree 2. This proves the lemma.

We let $\mathcal{P}_{\text{main}} := \{F \in \mathcal{P} : |V_F| = k + 1\}$. We also let

$$\mathcal{P}'_0 := \{F \in \mathcal{P} : |V_F| = k, F_P = F_Q, F(1, 0) = F(2, 0) = F(1, k) = F(2, k)\}.$$

Then, the contribution of $\mathcal{P}_{\text{main}} \cup \mathcal{P}'_0$ to (56) is given by

$$\sum_{\substack{F(1,0)=F(2,0) \in [n] \\ F(1,k)=F(2,k) \in [n]}} |u_{F(1,0)}|^2 |u_{F(1,k)}|^2 n^{k-1} = n^{k-1}.$$

We partition the remainder of \mathcal{P} in the following way. First let

$$T_1 := \{(1, 0), (2, 0), (1, k), (2, k)\} \subset T$$

be the terms corresponding to the starts and ends of the paths. For $t \geq 0$ and P a partition T_1 with $|P| \geq 2$ if $t = 0$, let

$$\mathcal{P}_{P,t} := \{F \in \mathcal{P} : |V_F| = k - t, F(s) = F(t) \Leftrightarrow s \sim_P t, s, t \in T_1\}.$$

Note that we exclude the trivial partition $P = \{T_1\}$ when $t = 0$ since $\mathcal{P}_{\{T_1\},0} = \mathcal{P}'_0$. We let $\mathcal{P}_0 = \bigcup_{P \neq \{T_1\}} \mathcal{P}_{P,0}$ and for $t > 0$, we let $\mathcal{P}_t = \bigcup_P \mathcal{P}_{P,t}$.

Lemma 11. For $F \in \mathcal{P}_t$, $W_x(F) \ll K^{2t}$.

Since $\mathbb{E}|x|^a \ll K^{(a-4)_+}$,

$$\begin{aligned} W_x(F) &\leq \prod_{e \in \mathcal{E}_F} \mathbb{E}|x|^{|e|} \\ &\ll \prod_{e \in \mathcal{E}_F} K^{(|e|-4)_+}. \end{aligned}$$

It suffices to show that $\sum_{e \in \mathcal{E}_F} (|e| - 4)_+ \leq 2t$. Since at most one vertex has no outgoing edge, $|\mathcal{E}_F| \geq k - t - 1$. Also the $|e|$'s satisfy $\sum_{e \in \mathcal{E}_F} |e| = 2k$ and $|e| \geq 2$. If $|e| \leq 4$ for all $e \in \mathcal{E}_F$, there is nothing to prove. If $|e_1| \geq 4$ say, then

$$\begin{aligned} \sum_{e \in \mathcal{E}_F} (|e| - 4)_+ &= |e_1| - 4 + \sum_{e \neq e_1} (|e| - 4)_+ \\ &\leq \sum_{e \in \mathcal{E}_F} (|e| - 2) - 2 \\ &\leq 2k - 2(k - t - 1) - 2 = 2t. \end{aligned}$$

We now turn to controlling $S_{P,t} := |\sum_{F \in \mathcal{P}_{P,t}} W_{u,v}(F)|$. To simplify notation, we will do this for the specific case $P = \{(1,0)\}, \{(2,0)\}, \{(1,k), (2,k)\}$. We can bound $S_{P,t}$ by

$$\sum_{\substack{i_1, i_2, i_3 \in [n] \\ \text{distinct}}} |u_{i_1} u_{i_2} v_{i_3}^2| |\{F \in \mathcal{P}_t : F(1,0) = i_1, F(2,0) = i_2, F(1,k) = F(2,k) = i_3\}|.$$

The cardinality of the last set is independent of the choice of indices i_1, i_2 and i_3 , and in fact only depends on size of the partition P . We denote it by $N_{|P|}$. Removing the restriction to distinct indices and using $\sum_i |u_i| = O(\sqrt{n})$, we may bound the contribution as $nN_{|P|}$.

The case for a general partition is similar and we have the bound

$$S_{P,t} \leq n^{c_P/2} N_{|P|}$$

where c_P is the number of singletons in the partition P . To determine $N_{|P|}$, we first choose the remaining vertices of V_F in $\binom{n}{k-t-|P|}$ ways. We let $N_2 = N_2(t)$ be the maximum number of ways to choose E_F , over P and V_F . Similarly, we let $N_3 = N_3(t)$ be the maximum number of ways to choose \mathcal{E}_F , over P , V_F and E_F . Since

$$\binom{n}{k-t-|P|} \leq \frac{n^{k-t-|P|} k^{|P|}}{(k-t)!},$$

we have

$$S_{P,t} \leq (n^{c_P/2-|P|} k^{|P|}) \frac{n^{k-t}}{(k-t)!} N_2(t) N_3(t),$$

with $|P| \geq 2$ if $t = 0$. Considering the possibilities for P and setting

$$S_t := \sum_{P \text{ partition of } T_1} S_{P,t},$$

we have

$$(57) \quad S_t \ll \begin{cases} k^2 n^{k-3/2} N_2 N_3 / k! & : t = 0 \\ kn^{k-t-1} N_2 N_3 / (k-t)! & : t \geq 1 \end{cases}.$$

We now estimate $N_2 = N_2(t)$, the number of ways to choose the set of edges E_F for $F \in \mathcal{P}_t$. As observed earlier, at least $k - t - 1$ vertices have positive outdegree, and similarly for the indegree. We need to assign at most k oriented edges to the $k - t$ vertices such that these conditions are met. Recall $d_{\text{out}}(i)$ to be the outdegree of vertex i . We will allow for repetitions when choosing the edges to include graphs with less than k edges. Hence we may impose the constraint $\sum_{i=1}^{k-t} d_{\text{out}}(i) = k$. For at least $k - t - 1$ vertices, $d_{\text{out}}(i) \geq 1$. This gives¹ $\binom{k}{t+1}$ ways of choosing the outdegrees $(d_{\text{out}}(i))_{i=1}^{k-t}$. To assign the incoming edges of the vertices, we partition the k edges into $k - t$ nonempty parts $(E_i)_{i=1}^{k-t}$. We first choose $k - t$ edges to belong to the different E_i 's and then we choose parts for each of the remaining edges. This can be done in at most $\binom{k}{t}(k - t)^t$ ways. Finally, we assign the $k - t$ parts to the vertices with positive indegree. If all $k - t$ vertices have incoming edges, there are at most $(k - t)!$ ways to assign each of them an E_i . Now suppose only $k - t - 1$ of the vertices have incoming edges. First, there are at most $(k - t)^4$ ways to choose 2 vertices and 2 parts, with one vertex being assigned both parts and the other having no incoming edges. Next, there are $(k - t - 2)!$ ways of assigning the remaining parts to the remaining vertices. Hence

$$\begin{aligned}
 N_2 &\leq \binom{k}{t+1} \binom{k}{t} (k - t)^t ((k - t)! + (k - t)^4 (k - t - 2)!) \\
 &\leq \frac{k^{2t+1}}{(t+1)!t!} k^t (k - t)! k^2 \\
 (58) \quad &\leq \frac{k^{3t+3} (k - t)!}{(t+1)!t!}.
 \end{aligned}$$

We now estimate $N_3 = N_3(t)$, the number of ways of choosing \mathcal{E}_F once V_F and E_F have been chosen. Since each vertex has at least one outgoing edge, the maximum outdegree of any vertex is at most $t + 1$. On the other hand, since $d_{\text{out}}(1) + \dots + d_{\text{out}}(k - t) \leq k$, at least $\max(k - 2t, 0)$ vertices have $d_{\text{out}}(i) = 1$. At least $\max(2k - 4t, 0)$ legs start from these vertices so at most $4t$ legs begin at vertices with $d_{\text{out}}(i) > 1$. At each of these legs, we have at most $t + 1$ choices to make when choosing the path. We thus have

$$(59) \quad N_3 \leq (t + 1)^{4t},$$

which is independent of the chosen vertices and edges.

For $t = 0$, using (57), (58) and (59), we have

$$\begin{aligned}
 S_0 &\leq n^{k-3/2} k^2 / k! N_2(0) N_3(0) \\
 &\leq k^5 n^{k-3/2}.
 \end{aligned}$$

Since $W_x(F) = O(1)$ for $t = 0$, the contribution to (56) is $o(n^{k-1})$.

¹This follows from the standard *stars and bars* combinatorial argument; see [9].

For $t \geq 1$, we have

$$\begin{aligned} S_t K^{2t} &\ll n^{k-t-1} k N_2 N_3 K^{2t} / (k-t)! \\ &\leq k^4 n^{k-1} \frac{k^{3t} (t+1)^{4t}}{n^{\epsilon t} (t+1)!} \\ &\leq k^4 n^{k-1} \frac{k^{3t} (te)^{2t}}{n^{\epsilon t}}, \end{aligned}$$

where we have used the estimates $t! > \frac{t^t}{e^t}$ and $\frac{(t+1)^t}{t^t} \leq e$. For $k = o(n^{\epsilon/5})$, the last expression is decreasing for $t \leq k$ and bounding each term by the bound for the $t = 1$ term, we have

$$\sum_t S_t K^{2t} \ll k^7 n^{k-1-\epsilon} = o(n^{k-1})$$

for $k = o(n^{\epsilon/7})$.

□

4. PROOF OF THEOREM 4

Proof. We will work on the event

$$E = E_n = \{\rho(X) < 1 + \epsilon, \lambda > 1 + 2\epsilon \text{ for all } \lambda \in \bigcup_{\theta \in \Theta} \Lambda^\theta\}$$

which occurs w.h.p. Fix $\theta \in \Theta$ and for $\lambda \in \Lambda^\theta$, let

$$R_\lambda := \left(\frac{X}{\sqrt{n}} - \lambda \right)^{-1}$$

denote the resolvent of X/\sqrt{n} . On E , $\lambda > \rho(X)$, so we may expand R_λ as a Neumann series

$$\begin{aligned} R_\lambda &= -\frac{1}{\lambda} \left(1 + \frac{1}{\sqrt{n}} \sum_{i \geq 1} \frac{X^i}{n^{(i-1)/2} \lambda^i} \right) \\ &=: -\frac{1}{\lambda} \left(1 + \frac{1}{\sqrt{n}} S_\lambda \right). \end{aligned}$$

We write the Jordan decomposition of A as $A = V J U^*$ where V (resp. U^*) is the $n \times \text{rk}(A)$ (resp. $\text{rk}(A) \times n$) matrix of generalized right (resp. left) eigenvectors of A associated to nonzero eigenvalues of A satisfying $U^* V = 1$ and J is the Jordan matrix of A restricted to nonzero eigenvalues with size $\text{rk}(A) \times \text{rk}(A)$. Starting with the eigenvalue equation $\det(\frac{X}{\sqrt{n}} + A - \lambda) = 0$ and using the determinant identity $\det(1 + AB) = \det(1 + BA)$, we have

$$\begin{aligned} \det \left(\frac{X}{\sqrt{n}} + A - \lambda \right) &= 0 \Rightarrow \det(1 + R_\lambda A) = 0 \\ &\Rightarrow \det \left(1 - \frac{1}{\lambda} U^* \left(1 + \frac{1}{\sqrt{n}} S_\lambda \right) V J \right) = 0 \\ &\Rightarrow \det \left(-\lambda + J + \frac{1}{\sqrt{n}} U^* S_\lambda V J \right) = 0. \end{aligned}$$

Let J_θ be the block matrix of J corresponding to eigenvalue θ and let U_θ^* and V_θ be the restrictions of U^* and V to the generalized left and right eigenvectors of θ respectively. Recall Proposition 2 as well as the notation used therein. We apply Proposition 2 with $M = J_\theta$ and $P = P^\theta = \frac{1}{\sqrt{n}}U_\theta^*S_\lambda V_\theta J_\theta$.

First note that for each column indexed by $t \in I_v^\theta$, $J_\theta e_t = \theta e_t$, where e_t is the coordinate vector corresponding to t . Hence for $s \in I_u^\theta$ and $t \in I_v^\theta$,

$$P_{st}^\theta = \frac{1}{\sqrt{n}}\theta u_s^* S_\lambda v_t.$$

Observe that the moment assumption made in Theorem 4 guarantees the applicability of Lemma 7 to the collection

$$\{\sqrt{n}P_{st}^\theta : s \in I_u^\theta, t \in I_v^\theta, \theta \in \Theta\}.$$

By Lemma 7, $(\sqrt{n}P_{st}^\theta)_{s,t,\theta} \Rightarrow (F_r)_{r \in I_2}$ defined by (8), (4) and (9). Finally, applying Proposition 2 yields the procedure to determine the fluctuations as specified in Theorem 4. \square

APPENDIX A.

In this section we state the deterministic perturbation result referred to in the proof of Theorem 4. It is originally attributed to Lidskii. See [14] and references cited within. We remind the reader that the Schur complement of A in the block matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is $D - CA^{-1}B$.

Proposition 2. *Let M be a $d \times d$ deterministic matrix in Jordan form. For notational simplicity, we will assume M has a single eigenvalue θ . Let*

$$J_k := \begin{bmatrix} \theta & 1 & & \\ & \theta & 1 & \\ & & \ddots & \ddots \\ & & & \theta \end{bmatrix}$$

denote the $k \times k$ Jordan block and write

$$M = \bigoplus_{k=1}^K J_k^{\oplus m_k}$$

Hence for each $k \in [K]$, M has m_k Jordan blocks J_k . Let P_n be a sequence of $d \times d$ perturbation matrices with entries of size $o(1)$. Then $M + P_n$ has spectrum

$$\Lambda(M + P_n) = \{\lambda_{k,m,i} : k \in [K], m \in [m_k], i \in [k]\}$$

with $\lambda_{k,m,i} \rightarrow \theta$ for all $k \in [K]$, $m \in [m_k]$ and $i \in [k]$. The fluctuations

$$f_{k,m,i} := \lambda_{k,m,i} - \theta$$

are given by the following procedure.

Let $c_k := \sum_{j=1}^k m_j$ and set $c := c_K$. Decompose $P = P_n$ into c^2 blocks $(B_{ij})_{i,j=1}^c$ with the c diagonal blocks $(B_{i,i})_{i=1}^c$ having sizes

$$1, \dots, 1, 2, \dots, 2, \dots, K, \dots, K$$

with k occurring with multiplicity m_k . Let $k_i \times k_j$ denote the size of block $B_{i,j}$. This block decomposition is conformal with that of M induced by the J_k 's. Let $R = R_n$ be the submatrix of P of size $c \times c$ with entries given by

$$R_{ij} = (B_{ij})_{k_i 1}.$$

Hence R is formed from the lower left elements of the blocks in the decomposition of P .

Let $E_k = R_{c_k \times c_k}$ be upper left submatrices of R and let F_k be the $m_k \times m_k$ Schur complement of E_{k-1} in E_k , where we set $F_1 := E_1$. Then, to leading order, the fluctuations $f_{k,m,i}$ are given by the k k -th roots of the m_k eigenvalues of F_k for each $k \in [K]$. If M has multiple eigenvalues, we apply the above procedure to each eigenvalue separately.

We remark on a few special cases of Proposition 2. We denote the entries of $P = P_n$ by p_{ij} and assume $p = O(\frac{1}{\sqrt{n}})$ (as will turn out to be the case in our applications).

- (1) Suppose $M = \text{diag}(\theta_1, \dots, \theta_d)$ is diagonal with distinct eigenvalues. Let λ_j denote the corresponding eigenvalues of $M + P$ in the sense that $\lambda_j \rightarrow \theta_j$ as $n \rightarrow \infty$. Then

$$f_j := \lambda_j - \theta_j = p_{jj}(1 + o(1)).$$

- (2) Suppose $M = \theta I_d$. Then $\{\sqrt{n}(\lambda_j - \theta)\}_{j=1}^d$ converge to the d eigenvalues of $\sqrt{n}P$.
- (3) Suppose $M = J_d(\theta)$. Then $\{n^{\frac{1}{2d}}(\lambda_j - \theta)\}_{j=1}^d$ converge to the d roots of $\sqrt{n}P_{k1}$.

APPENDIX B.

In this appendix, we extend the results involving the moment method, namely Proposition 1 and Lemma 7 using a truncation argument (see [1]). Consider the following two assumptions on the atom distribution x .

- (i) $|x| \leq K = O(n^M)$.
- (ii) $\mathbb{E}|x|^m < \infty$, $m = 2/M$.

We show that if Proposition 1 and Lemma 9 hold for (i) with $M < 1/2$, then they hold for (ii).

Suppose we have (ii) with $m > 4$, corresponding to $M = 2/m < 1/2$. We first show that the event

$$\{|x_{ij}| \leq n^M \text{ for all } i, j \in [n]\}$$

occurs w.h.p. Indeed, we have

$$(60) \quad \mathbb{P} [|x_{ij}| \geq n^M \text{ some } i, j \in [n]] \leq n^2 \mathbb{P} [|x|^m \geq n^2].$$

Since $n^2 \mathbb{1}_{|x|^m \geq n^2} \leq |x|^m$ and $\mathbb{E}|x|^m < \infty$, the last expression converges to 0 by the dominated convergence theorem.

Now define the truncated random variables $\hat{x} := x \mathbb{1}_{|x| \leq n^M}$ and $\hat{X} = (\hat{X})_{ij}$ by $\hat{X}_{ij} := \hat{x}_{ij}$. While \hat{x} is bounded, it no longer has mean zero. On the other hand, for n sufficiently large, we have

$$(61) \quad \begin{aligned} |\mathbb{E}\hat{x}| &\leq \mathbb{E}|x \mathbb{1}_{|x| \geq n^M}| \\ &\leq \frac{\mathbb{E}|x|^m \mathbb{1}_{|x| \geq n^M}}{n^{(m-1)M}} \\ &\ll n^{-(m-1)M} \\ &\leq n^{-3/2}. \end{aligned}$$

By Schur's test for the operator norm of a matrix, we have

$$(62) \quad \|\mathbb{E}\hat{X}\| = O(n^{-1/2}).$$

Now let $\tilde{x} := \hat{x} - \mathbb{E}\hat{x}$ and $\tilde{X} := \hat{X} - \mathbb{E}\hat{X}$ denote the truncated and centered random variables. By construction, $\mathbb{E}\tilde{x} = 0$. Furthermore,

$$(63) \quad \mathbb{E}|\tilde{x}|^2 = \mathbb{E}|\hat{x}|^2 - |\mathbb{E}\hat{x}|^2 \rightarrow \mathbb{E}|x|^2 = 1$$

by (61) and dominated convergence. Given (63), it is easy to check that under (i), Proposition 1 is valid for \tilde{X} . Since $\mathbb{E}|\tilde{x}|^2 \leq \mathbb{E}|x|^2$, Lemma 7 also valid for \tilde{x} . To prove the validity of Proposition 1 and Lemma 7 for x under (ii), it suffices to prove the following.

Lemma 12. *Suppose $u = u_n$ and $v = v_n$ are unit vectors in \mathbb{C}^n . Then for every $\gamma > 0$, the event*

$$A_{n,\gamma} := \bigcup_{k \leq \log^2 n} \{|u^* \hat{X}^k v - u^* \tilde{X}^k v| > \gamma n^{(k-1)/2}\}$$

occurs w.h.p.

We first state a result that is a consequence of the proof in [2]. Following the notation of [2] we define $\delta := n^{M-1/2}$ so that $|\tilde{x}| \leq \delta\sqrt{n}$. Fix $z > k+1$ and p a positive integer. Then

$$\begin{aligned} \mathbb{P} \left[\left\| \left(\frac{1}{\sqrt{n}} \tilde{X} \right)^k \right\| \geq z \right] &\leq z^{-2p} n^{-pk} \mathbb{E} \operatorname{Tr} \left(\tilde{X}^k \left(\tilde{X}^k \right)^* \right)^p \\ &=: z^{-2p} n^{-pk} E_n. \end{aligned}$$

In [2](pg. 561), it is shown that

$$E_n \leq n^{kp + \frac{3}{2}} \sum_{l=1}^{pk} \binom{2kp}{2l} (k+1)^{2kp-2l+2p} (2kp) \left(\frac{6kp\delta^{1/6}}{\log \frac{\delta\sqrt{n}}{(2kp)^3}} \right)^{6kp-6l} \delta^{kp-l}.$$

In our application, $k \leq \log^2 n$ and choosing $p = \delta^{-1/7}$ say, we have

$$(64) \quad \frac{6kp\delta^{1/6}}{\log \frac{\delta\sqrt{n}}{(2kp)^3}} \rightarrow 0.$$

In fact, the left-hand side of (64) is less than 1 for $n \geq N(m)$.

For such n , following [2](pg. 562), it then follows that

$$z^{-2p} n^{-pk} E_n \leq \left((2kpn^2)^{1/p} (1 + (k+1)\delta^{1/2})^{2k} \left(\frac{k+1}{z} \right)^2 \right)^p.$$

Choosing $z = 3k$ say, for any $k \leq \log^2 n$ we have

$$(65) \quad \mathbb{P} \left[\left\| \left(\frac{1}{\sqrt{n}} \tilde{X} \right)^k \right\| \geq 3k \right] = O(e^{-n^c})$$

for some $c = c(m) > 0$. We now turn to the proof of Lemma 12.

Proof. By (65), we may assume $\left\| \left(\frac{1}{\sqrt{n}} \tilde{X} \right)^k \right\| \leq 3k$ for all $k \leq \log^2 n$ which occurs w.h.p. We will need the crude bound

$$(66) \quad \sum_{\substack{a_1 + \dots + a_k = n \\ a_i \geq 0}} \prod_{i=1}^k a_i \leq n^{2k}.$$

We then have

$$\begin{aligned} \frac{1}{n^{(k-1)/2}} |u^* \hat{X}^k v - u^* \tilde{X}^k v| &\leq n^{-(k-1)/2} \|(\tilde{X} + \mathbb{E}\hat{X})^k - \tilde{X}^k\| \\ &\leq \sum_{l=1}^k \frac{1}{n^{(l-1)/2}} \sum_{l'=0}^{l+1} \sum_{a_1 + \dots + a_{l'} = k-l} \prod_{i=1}^{l'} \left\| \left(\frac{\tilde{X}}{\sqrt{n}} \right)^{a_i} \right\| \|\mathbb{E}\tilde{X}\|^l \\ &\leq \sum_{k=1}^l \frac{1}{n^{l-1/2}} \sum_{l'=0}^{l+1} (3k^2)^{l'} \\ &\leq \sum_{l=1}^k \frac{l+1}{n^{(l-1)/2}} \left(\frac{3k^2}{\sqrt{n}} \right)^{l+1} = o(1), \end{aligned}$$

where we have used (66) and (62) in the third line.

□

REFERENCES

- [1] Z. D. Bai and J. Silverstein. *Spectral analysis of large dimensional random matrices*. Mathematics Monograph Series **2**. Science Press, Beijing, 2006.
- [2] Z. D. Bai and Y. Q. Yin. Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang. *Probab. Theory Relat. Fields*, 73:555–569, 1986.

- [3] J. Baik, G. Ben Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.
- [4] J. Baik and J. W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.*, 97(6):1382–1408, 2006.
- [5] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [6] C. Bordenave and M. Capitaine. Outlier eigenvalues for deformed i.i.d. random matrices. *arXiv*, math.PR(1403.6001v2), 2014.
- [7] M. Capitaine, C. Donati-Martin, and D. Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. *Ann. Probab.*, 37(1):1–47, 2009.
- [8] M. Capitaine, C. Donati-Martin, and D. Féral. Central limit theorems for eigenvalues of deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(1):107–133, 2012.
- [9] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol 1*. Wiley, 2nd ed edition, 1950.
- [10] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.
- [11] A. Knowles and J. Yin. The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.*, 66(11):1663–1750, 2013.
- [12] A. Knowles and J. Yin. The outliers of a deformed Wigner matrix. *Ann. Probab.*, 42(5):1980–2031, 2014.
- [13] V. B. Lidskii. On the theory of perturbations of nonselfadjoint operators. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 6(1):52–60, 1966.
- [14] J. Moro, J. V. Burke, and M. L. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *Siam J. Matrix Anal. Appl.*, 18(4):793–817, 1997.
- [15] S. O’Rourke and D. Renfrew. Low rank perturbations of large elliptic random matrices. *Electron. J. Probab.*, 19:no. 43, 65, 2014.
- [16] A. Pizzo, D. Renfrew, and A. Soshnikov. On finite rank deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(1):64–94, 2013.
- [17] D. Renfrew and A. Soshnikov. On finite rank deformations of Wigner matrices II: Delocalized perturbations. *Random Matrices Theory Appl.*, 2(1):1250015, 36, 2013.
- [18] J. Rochet and F. Benaych-Georges. Outliers in the single ring theorem. *arXiv*, math.PR(1308.3064v4), 2013.
- [19] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
- [20] T. Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. *Probab. Theory Relat. Fields*, 155(1-2):231–263, 2013.
- [21] T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, 38(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur.
- [22] M. I. Višik and L. A. Ljusternik. Solution of some perturbation problems in the case of matrices and self-adjoint or non-selfadjoint differential equations. I. *Russian Math. Surveys*, 15(3):1–73, 1960.

DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES CA 90095-1555

E-mail address: anandbr@math.ucla.edu